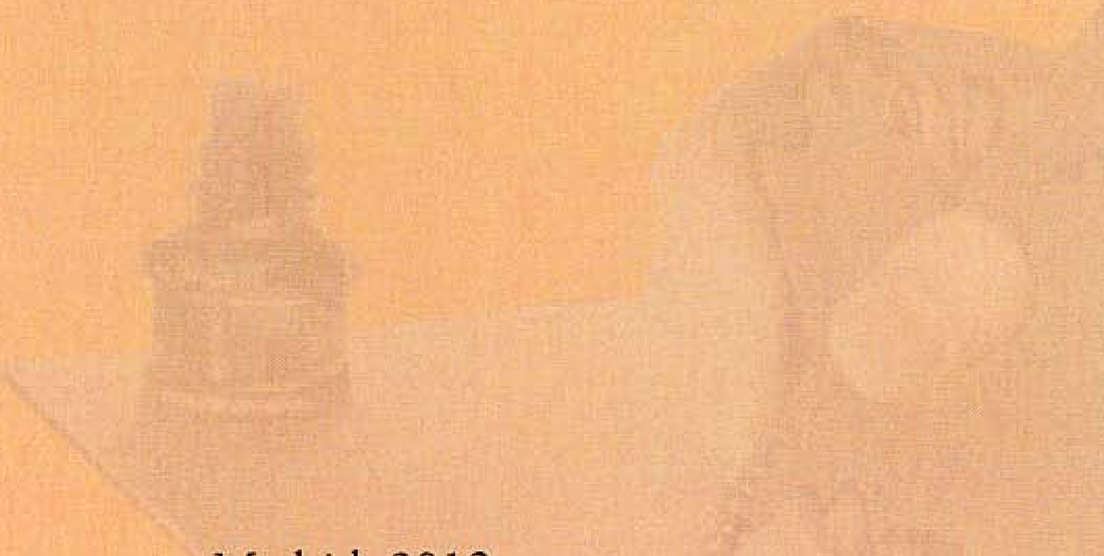
Isabel Del Bosque González
Carlos Fernández Freire
Lourdes Martín-Forero Morente
Esther Pérez Asensio

LOS SISTEMAS DE INFORMACIÓN GEOGRÁFICA
y la Investigación en Ciencias Humanas y Sociales



Madrid, 2012

Apuntes de Ciencias Instrumentales y Técnicas de Investigación. 3

Los Sistemas de
Información Geográfica
y la Investigación en
Ciencias Humanas y Sociales

ISBN: 978-84-615-9825-0
Dep. Legal: M-29450-2012
Edita: Confederación Española de Centros de Estudios Locales (CSIC)
Distribución y venta: CECEL-CSIC.
C/ Albasanz 26-28. 28037, Madrid.
cchs_cecel@cchs.csic.es

Los Sistemas de Información Geográfica y la Investigación en Ciencias Humanas y Sociales

Isabel del Bosque González
Carlos Fernández Freire
Lourdes Martín-Forero Morente
Esther Pérez Asensio

‘Cada individuo tiene su propio mapa del mundo. El del niño no se parece al del adulto…De ahí la dificultad de la comprensión mutua. Al hablar del mundo, cada cual tiene su propio mapa, su propia visión, su propia imagen”.

Ryszard Kapuscinsky (LAPIDARIUM IV, 2003)

Índice

Presentación

Casi todos los fenómenos, eventos u objetos de estudio de las Ciencias Humanas y Sociales se producen en un espacio geográfico determinado y pueden ser georreferenciados mediante un sistema de coordenadas espacio-temporal. En este marco, las tecnologías de información geográfica pueden mejorar notablemente la investigación científica en estas disciplinas abriendo otras perspectivas u oportunidades de conocimiento, planteando nuevas cuestiones, aplicando diferentes metodologías de análisis y revelando nuevos datos que sin el uso de estas tecnologías quizás permanecerían invisibles, en definitiva, facilitando la aparición de nuevas interpretaciones de la misma realidad. Su aplicación es relativamente reciente en el campo científico al que pertenecemos, aunque suficientemente probada en otras ciencias como las relacionadas con los recursos naturales, las problemáticas medioambientales o las ciencias agrarias, por ejemplo.

La importancia actual y futura de estas tecnologías junto a la existencia de nichos de desarrollo y el uso de las mismas en el Instituto de Economía, Geografía y Demografía o en el Instituto de Historia (y potencialmente en otros institutos) aconsejaron, ya desde los orígenes del Centro de Ciencias Humanas y Sociales (CCHS) del CSIC, la creación de una Unidad especializada de Sistemas de Información Geográfica (SIG) en nuestro Centro, que actuara de motor transversal y soporte técnico para las líneas de investigación implementadas en el CCHS, a

través de un elemento común: el territorio. De este modo se facilitan las tareas de diseño y gestión de bases de datos geográficas, el análisis espacial, la geovisualización, la representación cartográfica y la publicación, a través de las denominadas Infraestructuras de Datos Espaciales (IDE), de información científica espacialmente referenciada.

Proyectos del CCHS, como los que se mencionan en esta obra, en el ámbito de la Arqueología, la Historia, la Lengua, la Demografía o las Ciencias Políticas han ido incorporando en los últimos años, de forma novedosa en muchas ocasiones, estas técnicas de investigación basadas en los SIG y aplicando un enfoque de análisis básicamente espacial, favoreciendo un marco de trabajo integrador que facilita la generación de conocimiento científico, a la vez que posibilita la interdisciplinaridad y la creación de sinergias en la investigación en Ciencias Humanas y Sociales.

Eduardo Manzano Moreno
Director del
Centro de Ciencias Humanas y Sociales
(CCHS) del CSIC

Agradecimientos

La realización de este trabajo no habría sido posible sin la ayuda y paciencia de nuestros compañeros de la Unidad de Sistemas de Información Geográfica (SIG) del Centro de Ciencias Humanas y Sociales (CCHS) del Consejo Superior de Investigaciones Científicas (CSIC), tan responsables de los trabajos que aquí se exponen como nosotros mismos. Es de justicia por lo tanto destinar estas líneas de agradecimiento a Alfredo Gómez, Lorenzo Mateos, Victoria González, Rocío Gutiérrez, Roberto Maestre y Sara García, además de otros compañeros que han trabajado en la Unidad y colaborado con nosotros en distintos momentos y proyectos.

En el ámbito concreto de las Ciencias Humanas y Sociales estamos comprobando, con agrado, una demanda creciente en el empleo de estas tecnologías en investigaciones de naturaleza muy diversa. Por eso, queremos agradecer también a quienes desde el CCHS han confiado en las tecnologías de información geográfica y en la Unidad SIG, creando vías efectivas de integración de los SIG con sus propias disciplinas de estudio. Tal es el caso de Juan M. Vicent, con su dilatada trayectoria en este sentido y su aportación a lo que será la futura Infraestructura de Datos Espaciales (IDE) de Arqueología, de Javier Sanz, que dirige el proyecto EXTERSIAL, de Ana Crespo, líder del proyecto de investigación histórica DynCoopNet, de Pilar García, coordinadora en el proyecto del Atlas Lingüístico de la Península Ibérica (ALPI), de Diego

Ramiro, con su implicación en la IDE histórica de la ciudad de Madrid, Antonio Abellán y Clara Bécares, del proyecto SIGMayores, o de Antonio Uriarte y Alfonso Fraguas, por su colaboración en los proyectos arqueológicos.

Quisiéramos retroceder en el tiempo y remontarnos al origen de la actual Unidad SIG para agradecer a personas como Ascensión Calatrava, Directora del antiguo Instituto de Economía y Geografía, o Jesús González, Gerente del mismo, su apoyo en aquellos momentos. Por supuesto, agradecer a Felipe Criado, ex-Coordinador del Área de Humanidades y Ciencias Sociales, su apuesta e impulso efectivo en los inicios de la Unidad y a Ángel L. Rodríguez, de la Secretaría General Adjunta de Informática, ambos piezas fundamentales en el acuerdo tecnológico que el CSIC firmó con empresas punteras de software de Cartografía, SIG y Teledetección, y que hoy por hoy, posibilita el desarrollo y la materialización de gran número de proyectos en el CSIC.

Por último, nuestro agradecimiento sincero a la dirección del CCHS, muy especialmente a su Director, Eduardo Manzano, a los anteriores vicedirectores, Vicente Rodríguez y Manuel Molina y a los actuales, especialmente al Vicedirector Técnico, Diego Ramiro y al Gerente, Miguel Ángel López, por su incondicional apoyo en cada una de las iniciativas que desde la Unidad SIG proponemos.

Introducción

Durante décadas, los Sistemas de Información Geográfica (SIG) se han aplicado a problemas de gestión territorial y de recursos naturales, a cuestiones relacionados con el medioambiente, la logística militar o en contextos directamente vinculados con las ciencias de la Tierra, como la geografía, la geología, etc. Sólo recientemente se ha empezado a considerar el uso potencial de los SIG para otros campos y disciplinas relativamente inéditos[1,2] y en particular en la investigación en Ciencias Humanas y Sociales**.**

Los SIG han supuesto un cambio paradigmático *tecnológico* e *intelectual*[3], fundamentalmente en el ámbito de las geociencias[4] y de la cartografía, este paradigma debe ser entendido como el conjunto de procedimientos técnicos y metodológicos que permiten: por un lado, tratar la espacialidad de los datos, y por otro, favorecer el estudio de la realidad desde enfoques multidimensionales e integrados, como son el tiempo, el espacio y las "personas" que interactúan con el territorio

1. Anne Kelly Knowles y Amy Hillier, *Placing history: how maps, spatial data, and GIS are changing historical scholarship.* Redlands, Calif.: ESRI Press, 2008.
2. Ian Gregory y Paul S. Ell, *Historical GIS: technologies, methodologies, and scholarship.* Cambridge; New York: Cambridge University Press, 2007.
3. Gustavo D. Buzai y Claudia A. Baxendale, *Análisis socioespacial con Sistemas de Información Geográfica.* Buenos Aires: Lugar Editorial, 2006.
4. Las Ciencias de la Tierra o Geociencias son el conjunto de disciplinas que estudian la estructura, morfología, evolución y dinámica del planeta Tierra. http://es.wikipedia.org/wiki/Ciencias_de_la_Tierra.

en un momento determinado. Cuando las fuentes de información están debidamente vinculadas se posibilita la mejor compresión de los temas específicos y objetos de estudio[5].

Cualquier evento u objeto que pueda observarse en el territorio, puede ser representado cartográficamente referido a unas coordenadas espaciales y temporales. A menudo, el análisis y las formas de visualización de las relaciones espaciales de los objetos, añaden información que no sería visible de otra manera: "*explicita lo implícito de los datos*". En este sentido, cabe señalar que la tecnología SIG permite la modelización matemática de las relaciones espaciales, incluyendo el análisis de redes, la regresión espacial, la determinación de caminos óptimos y otras formas de análisis espacio-temporales que pueden ser muy valiosas para los estudiosos de la Arqueología, la Antropología, el Arte, la Historia, la Lengua, la Música y otras Ciencias Humanas, por no mencionar las Ciencias Sociales; favoreciendo además la interdisciplinaridad en la investigación científica y posibilitando nuevas formas de trabajo en contextos más tecnificados e innovadores, en unas disciplinas tradicionalmente muy compartimentadas y acostumbradas a trabajar en áreas muy especializadas.

Por ello, el propósito de esta obra es adentrarse en la teoría y los fundamentos de los SIG desde la perspectiva de su aplicabilidad a las Ciencias Humanas y Sociales, intentando enfatizar en los temas más relevantes para estas disciplinas e incidir en los conceptos fundamentales, proporcionando a los investigadores las referencias necesarias y el substrato teórico y técnico básico que les capacite para tratar con la componente espacial en su temática concreta de estudio. En este libro se presentan, por tanto, los fundamentos necesarios para el diseño y el modelado de datos espaciales y para la generación de bases de datos georreferenciadas, se abordarán las funciones elementales de consulta y edición, análisis espacial y representación de la información por técnicas cartográficas, para terminar esbozando unas ideas muy generales en cuanto a la difusión y publicación de la información georreferenciada mediante tecnología SIG en Internet, a través de las denominadas Infra-

5. Jack B. Owens, 'Toward a Geographically-Integrated, Connected World History: Employing Geographic Information Systems (GIS)', *History Compass*, 5 (2007).

estructuras de Datos Espaciales (IDE). A pesar de su aparente amplitud, es importante destacar que se trata de unos apuntes sobre los SIG y la investigación, y como tal, su objetivo es proporcionar una comprensión detallada de lo que es un SIG, cómo funcionan y cómo su aplicabilidad puede incidir favorablemente en las prácticas cotidianas de los investigadores. Con este fin, se ha tenido especial cuidado para asegurar que el enfoque de esta obra sea sobre la importancia de los SIG en las Ciencias Humanas y Sociales.

Ejemplos de SIG en Ciencias Humanas y Sociales.

No podemos hablar de los SIG en Humanidades y Ciencias Sociales sin empezar mencionando el *Center for Spatially Integrated Social Sciences (CSISS)*[6]. La integración de la perspectiva espacial en las teorías y prácticas de las Ciencias Sociales surgen, de hecho, a raíz de los trabajos desarrollados por este centro, donde la adopción de conceptos como distancia, ubicación, vecindad, proximidad, etc., y el uso de las tecnologías geoespaciales analíticas y la información geográficamente referenciada sirven de eje para abordar cuestiones fundamentales en estas disciplinas en el marco de la investigación, la gestión o la docencia[7].

La Arqueología y la Antropología son ciencias pioneras en la utilización de las tecnologías de información geográfica[8] (fundamentalmente los SIG y la Teledetección). De este modo, aplicados a estudios etnográficos, se encuentran algunos ejemplos de interés como los realizados en degradación de suelos y deforestación del Amazonas[9,10]. En cuanto a

6. CSISS fue creado en 1999 y financiado por la *National Science Foundation* americana (NSF BCS 9978058). Se encuentra ubicado en California, en la Universidad de Santa Bárbara. (*www.csiss.org*).

7. Michael F. Goodchild y Donald G. Janelle, *Spatially integrated social science.* Oxford: Oxford University Press, 2004.

8. David Wheatley y Mark Gillings, *Spatial technology and archaeology: the archaeological applications of GIS.* New York: Taylor & Francis, 2002.

9. Mark S. Aldenderfer y Herbert D. G. Maschner, *Anthropology, space, and geographic information systems.* New York: Oxford University Press, 1996.

10. Stephen D. McCracken et al., ‘Remote Sensing and GIS at Farm Property Level: Demography and Deforestation in the Brazilian Amazon’, *Photogrammetric Engineering and Remote Sensing*, 65 (1999) pp. (1311-1320).

los SIG y su aplicación a la Arqueología, la literatura científica es muy extensa[11,12], con casos de uso que van desde su utilización en la gestión de los recursos arqueológicos y la generación de grandes bases de datos geográficas, o el uso de las mismas en las propias excavaciones o en Arqueología del Paisaje, hasta su uso para generar diferentes tipos de simulaciones y modelos espaciales de análisis[13] de gran utilidad para arqueólogos en la interpretación del pasado, los cuales pueden ser combinados, además, con técnicas de geovisualización punteras e innovadoras como las de la realidad virtual o la modelización tridimensional[14].

Términos como "estadística espacial" o "econometría espacial" están empezando a utilizarse habitualmente en el campo de otras ciencias sociales como la economía[15], con aplicaciones sobre el crecimiento económico, las externalidades del conocimiento, la economía ambiental, el uso de la tierra, etc. Del mismo modo, para estudios de distribución de población, fertilidad, planificación familiar y otros temas relacionados con la demografía, la utilización de los datos espacialmente referenciados está abriendo nuevos campos de conocimiento[16],[17] donde, por ejemplo, la utilización de los SIG y la generación de bases de datos georreferenciadas a partir de los censos de población y vivienda son una línea de gran interés para el *Department of Economic and Social Affairs* de Naciones Unidas[18].

11. Gary R. Lock, *Beyond the map: Archaeology and spatial technologies.* Washington, DC: IOS Press, 2000.

12. Kathleen M. S. Allen, Stanton W. Green y Ezra B. W. Zubrow, *Interpreting space: GIS and archaeology.* London; New York: Taylor & Francis, 1990.

13. James Conolly y Mark Lake, *Sistemas de información geográfica aplicados a la arqueología.* Barcelona: Edicions Bellaterra, 2009.

14. Jun Liu y Guo-hua Geng, '3D Reality-based Modeling and Virtual Exhibition Technology for Culture Archaeological Sites', *Computer Engineering*, 2010 pp. (286-90).

15. Luc Anselin, Raymond J. G. M. Florax y Sergio J. Rey, *Advances in spatial econometrics: methodology, tools and applications.* Berlin (Germany): Springer-Verlag, 2004.

16. Chi-Chuan Chen, U. Sunday Tim y Sarah Stratton, 'Application of GIS in environmental epidemiology: assessment of progress and future trends', *GIS/LIS'96 Annual Conference and Exposition Proceedings|GIS/LIS'96 Annual Conference and Exposition Proceedings*, 1996 pp. (853-869).

17. G. William Skinner, Mark Henderson y Yuan Jianhua, 'China's fertility transition through regional space - Using GIS and census data for a spatial analysis of historical demography', *Social Science History*, 24 (2000) pp. (613-652).

18. United Nations. Statistical Division., *Handbook on geospatial infrastructure in support of census activities.* New York: United Nations, 2009.

En el contexto de la geografía humana se atribuye una importancia fundamental a la perspectiva espacial, tanto en los estudios de los seres humanos y su interacción con el medio que les rodea (*visión ecológica*) como en los de la diferenciación de áreas distintas sobre el territorio (*visión corológica*). Para cualquiera de estos dos enfoques, los SIG brindan la posibilidad de establecer correlaciones y asociaciones espaciales, favorecidas por la peculiar visión digital del mundo que proporcionan y sus posteriores tratamientos y análisis por medios informáticos[19]. Las aplicaciones son muchas y los casos de uso muy numerosos, desde estudios de paisaje, movimientos migratorios[20], temas relacionados con el cambio climático[21], geopolítica, planeamiento territorial, o en líneas más novedosas como las recientes aproximaciones a los SIG espacio-temporales para la representación tridimensional y el análisis de actividades humanas en espacios físicos y virtuales[22].

Otra aplicación interesante es la de los SIG como herramientas de apoyo a la toma de decisiones en el ámbito de las ciencias políticas[23]; en investigación en salud pública[24]; o en criminología, donde conceptos como la modelización espacial y los análisis de patrones que interrelacionan fenómenos sociales, económicos o ambientales con ocurrencias de delitos[25], están íntimamente ligados; o en sociología, con una larga tradición en la utilización de la componente geográfica[26], aquí la configuración espacial del territorio tiene especial importancia ya que afecta

19. Eric Sheppard, 'Quantitative geography: representations, practices, and possibilities', *Environment and Planning D-Society & Space*, 19 (2001) pp. (535-554).

20. David Robinson, 'The neighbourhood effects of new immigration', *Environment and Planning A*, 42 (2010) pp. (2451-2466).

21. Clionadh Raleigh y Henrik Urdal, 'Climate change, environmental degradation and armed conflict', *Political Geography*, 26 (2007) pp. (674-694).

22. Hongbo Yu y Shih-Lung Shaw, 'Exploring potential human activities in physical and virtual spaces: a spatio-temporal GIS approach', *International Journal of Geographical Information Science*, 22 (2008) pp. (409-430).

23. Christopher Thomas y Nancy Sappington, *GIS for decision support and public policy making.* Redlands, Calif.: ESRI Press, 2009.

24. Ellen K. Cromley, 'Spatial analysis, GIS, and remote sensing applications in the health sciences', *Professional Geographer*, 54 (2002) pp. (289-291).

25. Jacqueline Cohen, Wilpen L. Gorr y Andreas M. Olligschlaeger, 'Leading indicators and spatial interactions: A crime-forecasting model for proactive police deployment', *Geographical Analysis*, 39 (2007) pp. (105-127).

26. Ann R. Tickamyer, 'Space matters! Spatial inequality in future sociology', *Contemporary Sociology-a Journal of Reviews*, 29 (2000) pp. (805-813).

de forma muy decisiva en los comportamientos de los individuos y de los grupos que lo habitan y hace que las técnicas geoespaciales de modelado y análisis favorezcan, de manera muy eficiente, la comprensión de las conductas sociológicas y su investigación científica.

Cada vez es mayor el número de historiadores que se muestran interesados en el uso de los SIG[27,28] debido a la potencialidad que ofrecen las técnicas geoespaciales para el análisis y representación de los sistemas complejos —como los que tienen lugar en investigación histórica—, casi siempre dinámicos y a menudo no lineales[29], con la consiguiente demanda de herramientas que permitan organizar un gran número de variables e identificar las implicadas, como más probables, en la estabilidad y la transformación de tales sistemas. La combinación interdisciplinar entre SIG e investigación histórica es lo que se ha contextualizado, en términos generales, con el nombre de "SIG Históricos"[30].

Algunas de las manifestaciones originales de mayor envergadura han sido la creación de SIG históricos en países como Gran Bretaña[31] y Estados Unidos, ligados fundamentalmente a los límites administrativos, a la cartografía catastral y a la información de los censos antiguos.

Existen un gran número de preguntas en el ámbito de la investigación histórica: ¿qué pueden aportar los SIG en el contexto histórico?, ¿cómo se pueden utilizar los mapas antiguos en un marco analítico actual?, ¿qué papel juega la geovisualización en la transmisión de información a través del tiempo?; lo cierto es que las tecnologías geoespaciales permiten vincular el espacio geográfico con los acontecimientos históricos y moverse más allá de la cartografía estática para explorar las representaciones dinámicas, para pasar de sistemas pasivos a sistemas

27. Anne Kelly Knowles, *Past time, past place: GIS for history.* Redlands, Calif.: ESRI Press, 2002.

28. Ian Gregory y Paul S. Ell., *op. cit*.

29. Jack B. Owens et al., 'Visualizing Historical Narratives: Geographically-Integrated History and Dynamics GIS' presentado en Workshop on Visualizing the Past, 2009-02-20/ 2009-02-21., Richmond, Virginia (USA): University of Richmond, 2009.

30. Existe una red europea interesada en la aplicación de los SIG a la investigación histórica: " *The Historical GIS Research Network"*: http://www.hgis.org.uk/index.htm

31. Ian Gregory et al., 'Building a historical GIS for the British Isles', *AGI Conference at GIS 98. Profiting from Collaboration|AGI Conference at GIS 98. Profiting from Collaboration*, 1998 pp. (209-216).

interactivos, de las dos dimensiones a las representaciones multidimensionales, y trabajar en geovisualización, análisis exploratorio de datos geográficos y realidad virtual.

En el Centro de Ciencias Humanas y Sociales (CCHS) del Consejo Superior de Investigaciones Científicas (CSIC), se están llevando a cabo proyectos de investigación donde se están aplicando los SIG para abrir nuevos campos de conocimiento en estas disciplinas[32], algunos de los cuales se mencionan a continuación debido a su especial relevancia.

Caso de estudio 1: SIG e Historia

El proyecto DynCoopNet (*Dynamic Complexity of Cooperation-Based Self-Organizing Commercial Networks in the First Global Age*) forma parte del Programa EUROCORES, y fue aprobado tras su presentación a la convocatoria TECT (*"The evolution of cooperation and Trading"*) de la ESF (*European Science Foundation*)[33].

En el mismo participa un equipo internacional e interdisciplinar interesado en el estudio de la evolución de las redes de comercio auto-organizativas de los siglos XV al XVIII. El estudio introduce una novedad de carácter tecnológico, el uso de los SIG en la investigación histórica como motor de integración de datos y herramienta de visualización y análisis; el proyecto en sí supone un reto para las ciencias de la información geográfica por la dificultad que implica analizar y geovisualizar un sistema muy dinámico basado en las redes sociales colaborativas de gran complejidad, que dieron lugar a la primera economía mundial.

El SIG integra bases de datos históricas heterogéneas, dándolas significado global, con el objeto de establecer modelos de análisis his-

32. En la página web: http://humanidades.cchs.csic.es/cchs/sig/ existe información adicional de algunos proyectos SIG en Ciencias Humanas y Sociales.

33. Liderado por Ana Crespo Solana del IH (CCHS-CSIC). Referencias ESF: FP: 004DynCoopNet y Acciones Complementarias del Ministerio de Ciencia e Innovación: SEJ2007-29226. Este proyecto también está siendo financiado por el Programa MacroGrupos de la Comunidad de Madrid: Red de investigación: "Sólo Madrid es Corte. La construcción de la Corte de la Monarquía Católica", Programa MacroGrupos de la Comunidad de Madrid. Referencia CAM: HUM2007-29143-E/HIST.

tórico-comparativos (entre sistemas o procesos), econométricos (o de cuantificación) y de descripciones cualitativas de determinados fenómenos históricos[34]. Además, proporciona herramientas de geovisualización que ayudan a comprender la dinámica de este complejo sistema y la interacción social que dejó rastro en las fuentes documentales. En definitiva, permite realizar análisis encaminados a responder cuestiones acerca de la evolución de la cooperación de las redes comerciales autoorganizativas de la Primera Edad Global.

Algunas de las cuestiones que el SIG aborda son:

- Análisis del proceso evolutivo a lo largo del tiempo de las redes del sistema: modo de organización espacial de las redes colaborativas, modelos de cooperación entre agentes (relaciones de confianza, parentesco, reputación, proximidad geográfica, etc.).
- Identificación de las áreas del sistema que presentan mayor conectividad, e importancia de las diferentes zonas económicas en el conjunto del sistema (ej.: volumen comercial por puertos, distribución de agentes emisores-receptores en las relaciones epistolares,…).
- Análisis de mercados y su evolución histórica, comprobando sus diferencias espaciales, así como los métodos de las redes de mercaderes para asegurarse las vías de intercambio (rutas), negociar asientos o contratos, promocionar el oligopolio, o cómo funcionaban (en caso de poderse visualizar) los mercados de crédito y cómo se interconectaban geográficamente, estableciendo una jerarquía de las plazas financieras de forma paralela a la articulación de las redes mercantiles.
- Ejemplificar círculos de negocios con el fin de aplicar modelos matemáticos para modelar la naturaleza de los negocios (por ejemplo, el "viaje" espacio-temporal de una letra de cambio o de los contratos de seguros marítimos).

34. Esther Pérez et al., 'Integración de bases de datos históricas en una IDE. Comercio mundial y redes de cooperación en la primera Edad Global (1400-1800)', *VI Jornadas Técnicas sobre la Infraestructuras de Datos Espaciales de España (JIDEE 09)*, CD-Digital (2010) pp. (15).

- Análisis de la evolución de la percepción espacial e influencia de la cartografía en la emergencia de nuevas formas de cooperación humana.
- Comparaciones entre sistemas (Atlántico vs. Indopacífico) y visión del sistema único de la Primera Edad Global en conjunto. Estos sistemas podrían ser analizados y geovisualizados desde la perspectiva teórica y metodológica de los '*complex systems*' o método de las redes complejas, tradicionalmente empleado en otras disciplinas como la inteligencia artificial o la biología evolutiva, y que también se puede aplicar a la ciencia histórica desde el momento en que los procesos históricos analizados teleológicamente pueden considerarse sistemas complejos. Es el ejemplo de la formación de las denominadas "comunidades mercantiles", o la consolidación de lo que los historiadores llaman "economía-mundo", y que ha sido objeto de debate también en la geografía histórica. Ello servirá también para teorizar sobre qué es un sistema histórico.
- Comparación entre la actividad comercial generada por socios particulares (mercaderes) frente a la de los grandes sistemas de monopolio estatales.

Caso de estudio 2: SIG y Geolingüística

El Atlas Lingüístico de la Península Ibérica (ALPI) es un valioso documento en los archivos del CSIC, ya que supone un hito como atlas fonético de gran dominio peninsular, concebido por el equipo del filólogo e historiador Ramón Menéndez Pidal en el Centro de Estudios Históricos de la JAE y dirigido por su discípulo el lingüista español Tomás Navarro Tomás en los años treinta[35].

---

35. En el 2009, el CSIC retoma la publicación de la obra y lidera un macroproyecto –coordinado por Pilar García Mouton del ILLA (CCHS-CSIC)- que permita a la comunidad científica el acceso al ALPI completo, con todas las posibilidades de almacenamiento y búsqueda que las herramientas electrónicas permiten actualmente. Proyecto Intramural de referencia: 200410E604.

El SIG se está aplicando para la visualización de los datos lingüísticos en un entorno geográficamente referenciado, es decir, como nexo de conexión de la base de datos temática con una base de datos espacial, posibilitando no sólo el cartografiado automático a través de diferentes tipos de representaciones —centradas en la compartimentación del espacio mediante teselación voronoi o en la delimitación de unidades lingüísticas mediante isoglosas—, sino también la realización de búsquedas complejas en el corpus completo geolingüístico y etnográfico del ALPI.

Las tecnologías de información geográfica se usan también en este proyecto para generar una aplicación interactiva en Internet, a través de un geoportal con servicios Web basados en localización geográfica, de acuerdo a los estándares del *Open Geospatial Consortium* (OGC)[36] y combinados con búsquedas sobre la base de datos temática. La tecnología IDE, que apuesta por el acceso abierto a los datos y la implementación de servicios de acuerdo a los estándares del OGC, permitirá la interoperabilidad con otro tipo de información geográfica, aplicaciones y/o plataformas.

Lo novedoso del geoportal ALPI se refiere tanto a la componente temática —no hay referencias de IDE lingüísticas— como a la implementación de herramientas para la generación de cartografía temática *online* (mapas analíticos, léxicos, fonéticos y sintácticos) como resultado de la ejecución de consultas definidas por el propio usuario.

Caso de estudio 3: SIG y Arqueología

Durante años, un equipo de investigadores del CCHS–CSIC ha mantenido un proyecto interdisciplinar junto con colegas rusos en la zona de Kargaly (oblast de Orenburg), situada al sur de los Urales en la estepa rusa. El equipo español[37] se ha encargado de la caracterización

36. El *Open Geospatial Consortium* (OGC) es un consorcio internacional de más de 400 compañías, agencias gubernamentales y universidades que participan en un proceso de consenso para desarrollar estándares de interfaces de productos y servicios geográficos con criterios de interoperabilidad: http://www.opengeospatial.org/ogc.

37. La participación española ha sido dirigida por Mª Isabel Martínez Navarrete del IH (CCHS-CSIC) y financiada por los proyectos PS95-0031 (1996-1999), PB98-0653-(1999-2002) y BHA2003-08575(2003-2006).

paleoambiental de esta zona rica en mineral de cobre, objeto de una fuerte explotación en época prehistórica, entre el IV y II milenio cal. BC.

Los objetos de cobre procedentes de Kargaly tuvieron una gran difusión durante el período denominado Srubnaya, a finales de la Edad del Bronce, pero la evidencia arqueológica en cuanto a asentamientos y lugares funerarios no difiere sustancialmente de lo habitual para otros yacimientos de la misma cronología. Nos encontramos así ante una región en la que los trabajos mineros llevaron a la extracción de 150.000 toneladas de cobre[38], pero sin evidencias claras de su procesado.

Ante esta situación, los investigadores plantearon un acercamiento centrado en la Arqueología del Paisaje: en una zona mayoritariamente de estepa donde la principal fuente de combustible para los trabajos metalúrgicos se encuentra en bosques distribuidos en pequeñas manchas a lo largo de los ríos, las actividades metalúrgicas han de poder rastrearse en función de la evolución de dichos bosques.

De este modo, los SIG y la teledetección se convierten en la pieza central que articula los trabajos arqueológicos. La metodología desarrollada consistió en comparar la relación entre la lluvia polínica subactual y el paisaje moderno, con la lluvia polínica de la Edad del Bronce recogida a través de muestreos en turberas y zonas mineras, y datada por técnicas de carbono 14. Este acercamiento permitió evaluar la evolución histórica de las especies arbóreas más interesantes por su capacidad calorífica, como son el abedul y el roble, estableciendo una evolución diferencial provocada por la presencia de trabajos mineros en la zona norte de Kargaly frente a la zona sur[39], que muestra un desarrollo autónomo más relacionado con el impacto provocado por el pastoreo nómada.

38. Evgenij N. Chernykh, 'Kargaly: le plus grand ancien complexe minier et de métallurgie à la frontière de l´Europe et de l´Asie', ed por Marie-Chantal Frère-Sautot, Monographies instrumentumMontignac: Editions Monique Mergoil, 1998, v, pp. (71-76).

39. Juan M. Vicent García et al., 'Impacto medioambiental de la minería y la metalurgia del cobre durante la Edad del Bronce en Kargaly (región de Orenburgo, Rusia)', *Trabajos de Prehistoria*, 67 (2010) pp. (511-544).

Mediante un flujo de trabajo propio del ámbito de los SIG y la Teledetección, se generó el documento temático de cobertura del suelo necesario para cartografiar la presencia actual de las especies de interés y establecer la relación con la lluvia polínica en función de un mapa de distancias a bosques. De este modo, se pudo afirmar que en el entorno de los trabajos mineros existía un bosque de robles hoy desaparecido y que en general las muestras paleopolínicas recogían un porcentaje mayor de taxones relacionados con especies arbóreas que las muestras actuales, independientemente de su cercanía a los actuales bosques.

El proceso de trabajo requirió del concurso de otros métodos o tecnologías geoespaciales como el GPS (*Global Positioning System*) para georreferenciar las muestras de campo, el procesado de una imagen del sensor de teledetección ASTER para generar un Modelo Digital de Elevaciones (MDE), la clasificación digital de imagen satélite para obtener el mapa de coberturas del terreno o el cálculo de las distancias de todas las muestras a los bosques, entre otros.

Caso de estudio 4: SIG y Demografía

El proyecto de investigación desarrollado en el CCHS del CSIC pretende medir desde una perspectiva histórica un doble efecto sobre la demografía urbana[40]: el efecto de los flujos migratorios en una ciudad como Madrid, por un lado cómo los movimientos migratorios de cualquier índole pueden afectar a las variables demográficas (fecundidad, nupcialidad y mortalidad) en una gran ciudad y, por otro lado, el efecto de la afluencia de migrantes a instituciones de salud y caridad y cómo estas influyen en la mortalidad general y en otros indicadores demográficos de la capital. Para realizar los análisis, se ha escogido el período de años que va de 1896 a 1905, en el que se produjeron intensos movimientos migratorios hacia Madrid.

La información demográfica está vinculada a una base de datos cartográfica histórica adecuada ("*Plano de Madrid y pueblos colindan-*

40. Este es un proyecto de investigación del Plan Nacional I+D denominado "*La población de una gran ciudad: Madrid 1890-1935"* y liderado por Diego Ramiro Fariñas del IEGD (CCHS-CSIC). MICINN, CSO2008-06130/SOCI.

tes" de Facundo Cañada López del año 1902), de tal modo que permita representar las variables socio-demográficas de estudio a suficiente nivel de detalle, así como los posteriores trabajos de análisis geoespacial[41].

En un extremo del espectro, a través del SIG se puede representar toda una serie de variables demográficas a través de mapas temáticos (en indicadores como tasas de natalidad, nupcialidad, mortalidad por causas específicas, etc.) para unidades agregadas, como barrios o distritos. El nivel de detalle de la cartografía no obstante, permite el cálculo y la representación de esos mismos indicadores a nivel de manzana, permitiendo mapas mucho más completos de la diversidad demográfica madrileña. Dada la información padronal disponible, además, se pueden representar otros indicadores relacionados con la densidad, precio de las viviendas, ocupación, estatus socio-económico y demás variables asociadas a las manzanas urbanas.

Más allá de las explotaciones meramente descriptivas, se pueden realizar toda una serie de tipos de análisis geospacial sobre la distribución de los eventos o variables en la demografía madrileña, relacionadas con otros factores, así como análisis geoestadísticos, como por ejemplo la detección de concentraciones de unidades con características similares, pautas de segregación espacial, evolución de epidemias en la geografía urbana, etc.

Caso de estudio 5: SIG y Ciencia Política

El proyecto EXTERSIAL[42] tiene como principal objetivo el desarrollo de metodologías de análisis y valoración de intangibles territoriales, destinados a ser empleados en el diseño de políticas públicas, que integran enfoques tanto de oferta como de demanda de externalidades y

---

41. Isabel Del Bosque González et al., 'Cartografía y demografía histórica en una IDE. WMS del plano de Madrid de «Facundo Cañada»', *Revista Catalana de Geografía IV època*, 2010 <http://www.rcg.cat/articles.php?id=180> [accedido 26 Agosto 2011].

42. "*Externalidades territoriales en sistemas agroalimentarios locales: Desarrollo rural, paisajes y bienes públicos en denominaciones de origen de aceite de oliva".* El Investigador principal del mismo es Javier Sanz Cañada del IEGD (CCHS-CSIC). MICINN. Proyecto del Plan Nacional de I+D+i: CSO2009-08154.

se caracterizan por incorporar dos fuentes de innovación: i) tener como objeto de estudio los Sistemas Agroalimentarios Locales (SIAL), frente al predominio, casi exclusivo en la bibliografía, de los sistemas agrarios; ii) incorporar al análisis de las externalidades medioambientales, las externalidades vinculadas al capital institucional, al capital social o al patrimonio cultural, así como otros factores inherentes a la gobernanza territorial de los SIAL. El sector objeto de estudio es el aceite de oliva. Las dos zonas de trabajo escogidas para la realización del análisis empírico son las Denominaciones de Origen Protegidas de Estepa (Sevilla) y de Sierra de Segura (Jaén).

Desde el punto de vista de la *oferta de externalidades territoriales*, el proyecto se centra en el análisis de externalidades agroambientales, patrimonio natural y paisajes agrarios, así como de las externalidades de tipo socioeconómico y cultural.

La aportación de la tecnología SIG, para ello, es muy evidente:

- Caracterización del medio físico y su respectivo grado de adecuación para el cultivo del olivar. Con la generación de modelos digitales del terreno; geodatabases a partir de ortofotos, imágenes de satélite, mapas topográficos, geomorfológicos, edafológicos e información climática; generación de mapas de pendiente, de orientación del suelo, de índices y variables climatológicas, etc.

- Caracterización del medio natural: definición de mapas de usos y cubiertas del suelo, series de vegetación, indicadores de diversidad paisajística de los agroecosistemas, existencia de olivos singulares, parámetros e índices que evalúan el grado y el estado de la cubierta vegetal (índice de área foliar, el de vegetación de diferencia normalizada, el de radiación fotosintética activa, etc.).

- Caracterización de los agrosistemas de olivar: algunos indicadores se obtienen mediante cartografía, otros mediante información estadística georreferenciada y otros a partir de las encuestas a las explotaciones olivícolas.

Mediante técnicas SIG, también se obtienen diferentes tipos de *indicadores territoriales de síntesis:* grado de marginalidad física, mediante la capacidad de uso o aptitud del medio físico para el cultivo de olivar; grado de marginalidad económica, a partir del cálculo de la estructura de costes; si existe correlación espacial, y en qué grado, entre los fenómenos de marginalidad física y económica, potencialidad de erosión, colonización vegetal, configuración espacial del paisaje, análisis de las correlaciones espaciales e interrelaciones causales entre los indicadores de síntesis mediante procedimientos geoestadísticos, etc.

Caso de estudio 6: Geoportales y servicios Web basados en localizaciones geográficas

Aunque la aparición de las IDE es un fenómeno relativamente nuevo, la rápida proliferación de geoportales Web ha supuesto un *"punto de inflexión"* en la concepción de algunos de los principios fundamentales sobre los que se apoyaba la cartografía y la información geográfica hasta el momento, en cuanto a la adaptación de nuevas metodologías, sistemas organizativos y formas de trabajo en contextos innovadores, y un desarrollo, sin precedentes, de los mecanismos de difusión y utilización de los datos geoespaciales.

La creciente necesidad y el derecho de los ciudadanos al acceso a los datos y servicios geográficos de referencia, junto con iniciativas globalizadoras como la de *open access* y *open source* que posibilitan compartir información y desarrollos tecnológicos, unidos a la potencialidad de la Red en sí misma, y a importantes directivas europeas como INSPIRE[43] han permitido que las IDE se conviertan en importantes sistemas telemáticos para el desarrollo de la "*Sociedad de la Información"*, además de ser herramientas estratégicas en la formulación de un amplio abanico de políticas públicas.

INSPIRE propone un modelo de IDE como la solución tecnológica adecuada para compartir y publicar todo tipo de información georre-

43. Directiva Europea 2007/2/EC para el establecimiento de una ***IN**frastructure for **SP**atial **I**nfo**R**mation in the **E**uropean Community,* disponible en http://inspire.jrc.ec.europa.eu/

ferenciada en plataformas digitales, permitiendo a cualquier usuario el descubrimiento, acceso y uso generalizado a los datos geoespaciales, a través de Internet, por medio de motores de búsqueda como son los catálogos estandarizados de metadatos.

En el CCHS del CSIC se han abordado, de forma pionera, IDE científicas en el contexto de las Ciencias Humanas y Sociales. Se encuentran, totalmente operativas, la "*IDE sobre recursos sociales en España*" (http://www.sigmayores.csic.es), diseñada y desarrollada como apoyo a una de la líneas prioritarias del CSIC (envejecimiento y calidad de vida) y algunos de los geoportales de patrimonial cultural más significativos, como IDEZAM: *IDE de paisajes culturales de la zona de "Las Médulas"* (http://www.idezam.es/) o SILEX: *IDE del yacimiento arqueológico de "Casa Montero"* (disponible en http://www.casamontero.org); por otro lado, se encuentran en fase avanzada de desarrollo, una IDE sobre cartografía y demografía histórica de la ciudad de Madrid o la edición digital e interactiva del ALPI.

Capítulo I
Fundamentos de los Sistemas de Información Geográfica

Este capítulo está orientado a describir los conceptos y principios básicos de los SIG (ó GIS, frecuentemente se usa también su acrónimo en inglés). Se presentan las acepciones y definiciones más importantes, los componentes en los que se estructuran estos sistemas de información, se explorarán las formas complejas de las que se valen los SIG para modelizar el mundo real, su transformación en modelos virtuales y cómo se estructuran éstos geométricamente, para terminar con una breve introducción a los conceptos cartográficos básicos, necesarios para entender la obligada vinculación espacial entre los distintos objetos de la base de datos geográfica, sólo posible, si la información está georreferenciada en el mismo marco geodésico y en idéntico sistema de coordenadas.

1. ¿Qué es un SIG?: Algunas definiciones

Basta consultar los libros o textos más conocidos escritos sobre la materia, para comprobar que los distintos autores no muestran un consenso sobre lo que se debería entender por sistema de información geográfica, como se puede apreciar en la tabla 1. El motivo de ello es que existen, básicamente, tres perspectivas distintas de aproximación[44]

44. Peter A. Burrough y Rachael A. McDonnell, *Principles of Geographical Information Systems*, 2o edn. Oxford University Press, USA, 1998.

al concepto de SIG: una basada en las funcionalidades, en la capacidad para crear modelos digitales de la realidad y en la potencialidad de las herramientas que esta disciplina proporciona; otra entendida como una extensión del concepto en sí de base de datos, enfatizando en las diferencias en cuanto a la estructuración de los mismos y las distintas posibilidades de manejo de la información espacial; y la tercera atendiendo a aspectos que podríamos denominar "organizativos" o de las instituciones y personas que manejan la componente geográfica, donde los SIG se configuran como potentes sistemas de información de ayuda a la gestión y a la toma de decisiones.

Los datos geográficos o espaciales representan fenómenos reales o abstractos en términos de: a) su posición y forma respecto de un sistema de coordenadas conocido, b) los atributos (geoespaciales y temáticos o descriptivos) relacionados con el objeto geográfico, como podrían ser la elevación, temperatura, color, etc., y c) las interrelaciones espaciales existentes entre los objetos. Esto último es conocido con el nombre de "topología", un concepto clave en el mundo de los SIG que se refiere a la información cualitativa compuesta por un conjunto de reglas que describen las propiedades y conexiones, no métricas, de la información geográfica, como pueden ser la contigüidad, la conectividad, la superposición o la adyacencia. La topología es, además, necesariamente invariable frente a las transformaciones o las distorsiones geométricas que puedan sufrir los propios datos debidas, por ejemplo, a los cambios de proyección, escala o sistema de referencia.

La frontera entre los SIG y otros sistemas informatizados, como pueden ser los CAD (*Computer Assisted Design),* que permiten la representación cartográfica mediante la utilización de herramientas informáticas, los *mapping,* donde se incorporan, además, la posibilidad de trabajar con información georreferenciada, o con un sistema de Gestión de Bases de Datos (SGBD), es a veces bastante difusa. La mayoría de los autores coinciden en que lo más característico de un SIG es su capacidad de modelización del territorio, su potencial integrador de fuentes de información heterogéneas y diversas, y sobre todo su dimensión de análisis espacial[45], la posibilidad de realizar tareas complejas de

45. Joaquín Bosque Sendra, *Sistemas de Información Geográfica*, 2o edn. Rialp, 1997.

geoprocesamiento a partir de las relaciones topológicas de los objetos geográficos y, sobre todo, la capacidad de generar nueva información y/o explicitar la ya existente en su propia base de datos.

Tabla 1: Definiciones de Sistemas de Información Geográfica
"*Un potente conjunto de herramientas para recoger, almacenar, recuperar, transformar y visualizar datos geoespaciales del mundo real para un conjunto particular de propósitos"* (Burrough, P.A. and Mcdonnell, R., 1998).
"Un SIG es un sistema para capturar, almacenar, validar, integrar, manipular, analizar y representar datos referenciados sobre la Tierra" (Departamento de Medioambiente Británico, 1987).
"Un sistema compuesto de hardware, software y procedimientos para la captura, gestión, manipulación, análisis, modelado y representación de datos georreferenciados, con el objetivo de resolver problemas complejos de planificación y gestión" (National Center for Geographic Information and Analysis [NCGIA] of USA, 1990).
"Una tecnología de la información que almacena, analiza y visualiza datos espaciales y no espaciales" (Parker, 1988).
"Un conjunto de procedimientos manuales o informáticos utilizados para almacenar y manipular información georreferenciada" (Aronoff, 1989).
"Un sistema de bases de datos en el cual la mayoría de los mismos están espacialmente indexados, y que dispone de un conjunto de procedimientos destinados a responder cuestiones acerca de las entidades espaciales de la base de datos" (Smith et al., 1987).
"Una base de datos espacializada que contiene objetos geométricos" (Cebrián, 1994).
"Una entidad institucional reflejo de una estructura organizativa que integra tecnología con bases de datos, conocimiento experto y financiación continuada en el tiempo" (Carter, 1989).
"Un sistema de apoyo o ayuda a la toma de decisiones que integra datos espacialmente referenciados en un contexto de resolución de problemas" (Cowen, 1988).
"Un Sistema de Información Geográfica (SIG o GIS, en su acrónimo inglés [Geographic Information System]) es una integración organizada de hardware, software y datos geográficos diseñada para capturar, almacenar, manipular, analizar y desplegar en todas sus formas la información geográficamente referenciada con el fin de resolver problemas complejos de planificación y gestión. También puede definirse como un modelo de una parte de la realidad referido a un sistema de coordenadas terrestre y construido para satisfacer unas necesidades concretas de información. En el sentido más estricto, es cualquier sistema de información capaz de integrar, almacenar, editar, analizar, compartir y mostrar la información geográficamente referenciada. En un sentido más genérico, los SIG son herramientas que permiten a los usuarios crear consultas interactivas, analizar la información espacial, editar datos, mapas y presentar los resultados de todas estas operaciones". (http://es.wikipedia.org/wiki/Sistema_de_Información_Geográfica [29/06/2011]).

Especialmente oportuna es la distinción realizada por A. Rodríguez Pascual[46], que advierte de la confusión que puede generar, en contextos no especializados, el hecho de utilizar el mismo acrónimo (el término SIG) para referirse, conceptualmente, a cuestiones esencialmente distintas:

- SIG como *disciplina*: es el ámbito en el que se mueven las definiciones anteriores y lo expresado aquí hasta el momento.
- SIG como *proyecto*: cada una de las realizaciones prácticas o implementaciones técnicas de la disciplina SIG; así se pueden encontrar ejemplos que aluden a SIG arqueológicos, SIG medioambientales o el SIG de una u otra comunidad autónoma, etc.
- SIG como *software*: La industria tiende a referirse con el término SIG al programa o conjunto de programas que posibilitan desarrollar esta disciplina e implementar los proyectos correspondientes, de esta forma se habla de SIG propietarios, SIG libres o del SIG de una u otra casa comercial.

Diversos campos han contribuido a los fundamentos de los SIG[47]. La Geodesia, la Topografía y la Cartografía han permitido establecer las reglas y herramientas necesarias para medir y representar las distintas características del mundo real, aportando el sustrato teórico y técnico necesario que hace posible la georreferenciación de los datos digitales. La Informática y otras ciencias computacionales ofrecen el marco adecuado para el almacenamiento y manejo de la información geográfica y, junto con las Matemáticas, contribuyen con las herramientas necesarias para la manipulación de los objetos geométricos que representan la realidad-terreno. Datos de tipo socio-económico, ambiental, topográfico, etc., rellenan la gran variedad temática de la base de datos geográfica, de tal forma que, como ya sabemos, las aplicaciones SIG están presentes en una gran diversidad de áreas, desde las más académicas a las más comerciales.

46. Se puede consultar en los temas on-line de la oposición para el ingreso en el cuerpo nacional de ingenieros geógrafos, en la página Web del Instituto Geográfico Nacional: http://www.ign.es/ign/resources/acercaDe/aig/B.pdf [accedido 30 Enero 2011].

47. United Nations. Statistical Division, *op. cit*.

Por ello, algunos autores se refieren a esta disciplina como de "Ciencia de la Información Geográfica"[48]. El principal defensor de esta teoría es M. Goodchild[49], quien acuñó la idea y el término, en inglés "*GIScience",* en un artículo a principio de los noventa, refiriéndose al conjunto de conceptos y técnicas que subyacen en los SIG. El autor argumenta que existen ciertos aspectos que se plantean dentro de estos sistemas, de especial relevancia, y cuyo estudio sistemático constituye una ciencia por sí misma. Cuestiones como por ejemplo: ¿cómo puede un usuario conocer la precisión de los resultados?, ¿qué principios pueden aplicar los usuarios SIG para obtener mejores representaciones cartográficas?, ¿cómo pueden personalizarse las herramientas haciéndolas más compresibles para los usuarios no especializados?, etc., abren un campo de posibilidades dentro de la investigación científica. Al igual que las ciencias de la información plantean cuestiones fundamentales acerca de la creación, manejo, almacenamiento y uso de la información, *GIScience* puede dedicarse al estudio, desde el punto de vista científico, de la información geográfica[50]. Ciencia o tecnología, lo cierto es que cada vez hay un número mayor de revistas especializadas en torno a este tema, a la vez que emergen iniciativas globalizadoras y de cohesión muy importantes como *"The US University Consortiun for Geographic Information Science"* (http://www.ucgis.org), que agrupa aproximadamente a 60 universidades y asociaciones profesionales dedicadas al fomento y la promoción de la investigación en la ciencia de información geográfica, o la conferencia internacional *GIScience* (http://www.giscience.org), para la divulgación de los resultados de la investigación en todos los sectores de este campo y que se celebra bianualmente, de forma ininterrumpida, desde el año 2000.

2. Componentes de un SIG

Un SIG está integrado por un conjunto de subsistemas donde convergen aspectos muy tecnológicos con otros más tradicionales. Lo for-

48. Véase por ejemplo: A. Stewart Fotheringham y John P. Wilson, *The handbook of geographic information science.* Oxford: Blackwell, 2008.

49. Michael. F. Goodchild, 'Geographical Information science', *International Journal of Geographical Information Systems*, 6 (1992) pp. (31-45).

50. Paul A. Longley et al., *Geographic Information Systems and Science.* Wiley, 2001.

man una serie de elementos estrechamente relacionados entre sí, en los que cada uno de ellos posee identidad propia, pero el conjunto de todos genera una herramienta de potencial extraordinario.

La visión tradicional y la más recurrida en la bibliografía es la de dividir el SIG en cinco componentes básicos: hardware, software, datos o información geográfica, personas que interactúan con el sistema y procedimientos organizativos propios.

- *Hardware:* En realidad no hay un grupo de componentes informáticos físicos diferentes de otros sistemas que gestionen y manipulen grandes bases de datos gráficas, a excepción en todo caso, de los dispositivos de entrada de datos, como son los escáneres de planos o los de salida, como pueden ser los plotters.
- *Software:* Se refiere al conjunto de aplicaciones informáticas necesarias para la construcción, gestión, mantenimiento y explotación de la base de datos geográfica, ver Figura 1. Tanto los software SIG comerciales como los de código libre poseen, además de las funcionalidades básicas, herramientas que permiten crear mapas a medida del usuario y rutinas propias de programación "orientada a objetos", que posibilitan personalizar las aplicaciones, así como aplicaciones específicas para el desarrollo de servidores de mapas y cartografía en Internet, con las posibles descargas de datos, consultas y análisis de forma remota a través de dispositivos fijos o móviles.
- *Datos:* Representan la realidad geográfica y son el núcleo vital del SIG. Poseen unas características específicas en cuanto a la forma, y unas propiedades que le son inherentes que se precisan conocer en detalle para la adecuada rentabilidad del sistema.
- *Personas:* Hay distintas personas, con diferentes roles, que intervienen en un proyecto SIG. Desde los usuarios de diferentes disciplinas que utilizan el sistema para unas necesidades específicas, a los que marcan los requerimientos del mismo, pasando por el personal especializado tanto en tecnologías de la información como en las ciencias y técnicas afines con la Cartografía (Fotogrametría, Teledetección, Topografía, Geografía, etc.) implicado en tareas de diseño, desarrollo y adaptabilidad del SIG.

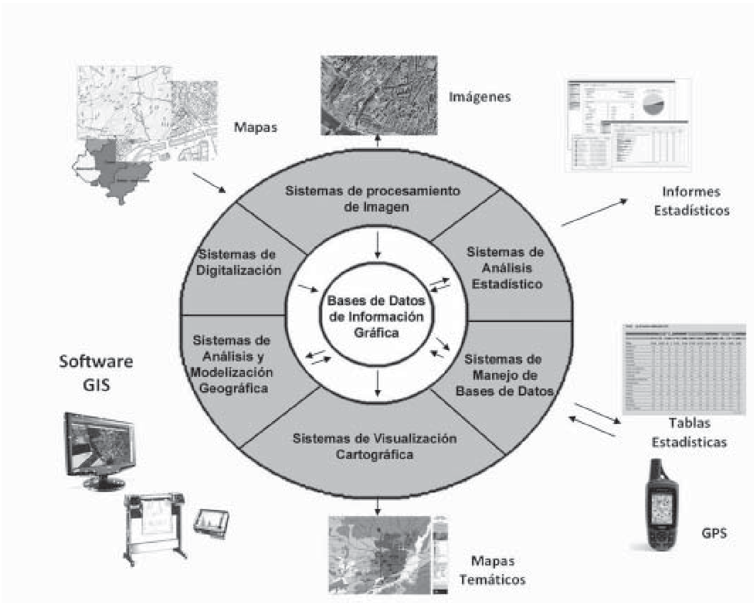


Figura 1. Funcionalidad básica de un SIG.Fuente: Adaptación de Ebener/WHO.

- *Organización:* El factor organizativo ha ido ganando importancia a medida que los SIG crecían en complejidad y se multiplicaban las aplicaciones prácticas. Una buena organización es fundamental para que el sistema funcione adecuadamente, donde se establezcan: definiciones claras de objetivos, procesos de planificación, coordinación, procedimientos normalizados de gestión enfocados a la funcionalidad, controles de calidad, etc.

Con la incorporación de nuevos profesionales y campos de aplicación al mundo de los SIG, y la natural evolución de los mismos en un contexto global y fuertemente interdisciplinar, hay autores que plantean la redefinición de algunos de los conceptos básicos de esta disciplina, adoptando una evolución del esquema clásico de cinco elementos con la incorporación de un sexto componente[51], cada uno de los cuales actuaría como un pilar conceptual sobre el que se apoya el estudio de la disciplina SIG[52]. De acuerdo a esta nueva perspectiva se introduce un

nuevo elemento de gran relevancia, que actúa de manera envolvente del resto que es la *geovisualización*.

Las técnicas cartográficas para representar la información han evolucionado de forma paralela a los SIG y a nadie se le escapa la trascendencia de una efectiva representación de los datos y el poder de la visualización de la información[53]. La cartografía del siglo XXI es dinámica y multitemporal, pero sobre todo es interactiva, permitiendo la representación del tiempo y del espacio adaptándose a las necesidades del usuario final, con formas transgresoras como la visualización tridimensional, la realidad aumentada o las representaciones gráficas SOM *(self-organizing map)* generadas a partir de redes neurales artificiales y que permiten preservar las propiedades topológicas del espacio. Es por ello que la geovisualización es un elemento poderoso como sistema de comunicación en sí mismo y que la forma en la que se representan los datos puede tener una influencia decisiva en la toma de decisiones y en la interpretación de resultados.

2.1. La Información Geográfica

Integrar la realidad geográfica en un sistema informatizado es una tarea ardua y delicada, que requiere un alto nivel de abstracción y generalización espacial. El mundo real es especialmente complejo y hace que sea casi imposible representarlo de manera exhaustiva en cualquier sistema. Es por ello que se hace absolutamente necesario la simplificación de la realidad y la conceptualización de la misma a través de modelos matemáticos.

Hay que realizar la distinción entre lo que se entiende por "datos", "información" y "conocimiento"[54]. Los datos son la representación concreta de la realidad, un conjunto de datos interrelacionados forman

51. Se puede consultar la página web: http://gislounge.com/the-components-of-gis-evolve/ [accedido 06 Julio 2011].

52. Víctor Olaya, *Sistemas de Información Geográfica*, licencia Creative Common Atribucion, 2010 <http://www.sextantegis.com/> [accedido 7 Junio 2011].

53. Jacques Bertin, *Semiology of Graphics: Diagrams, Networks, Maps*, 1o edn. ESRI Press, 2010.

una base de datos —de la cual se sobreentiende que está en formato digital—; la información es fruto del proceso interpretativo del usuario a partir de la base de datos, utilizando diferentes técnicas, medios o habilidades basadas en el conocimiento.

Los datos geográficos poseen unas características concretas que los hacen diferentes de otro tipo de datos, a la vez que los añade dificultad; estas peculiaridades son cuatro: posición, atributos descriptivos, relaciones espaciales y temporalidad[55]. Por otro lado, los fenómenos pueden expresarse geográficamente como "datos continuos", en los cuales existe un valor determinado para el fenómeno geográfico en cada punto del área de estudio (por ejemplo, un modelo digital del terreno) y como "datos discretos", donde el fenómeno geográfico se representa con límites claramente diferenciados.

El principio de organización de los datos, en el que se basan los SIG para poder caracterizar de forma adecuada el territorio, es el de estructurar la información en forma de capas temáticas vinculadas geográficamente, tal y como se muestra en la Figura 2a, donde cada una de las capas hace referencia a un tipo de datos o a un conjunto de los mismos de características similares. Esto permite al usuario combinar capas y crear nueva información a partir de ellas o realizar consultas complejas que involucren más de un tema. La capacidad de integrar datos de fuentes heterogéneas mediante el uso de la ubicación geográfica, o lo que es lo mismo, la utilización de la componente espacial como sistema de indexación es, de hecho, uno de los beneficios más importantes que pueden aportar los SIG. A este respecto, la FAO[56] afirmaba en el 2006: "*Un SIG almacena la información en capas temáticas que pueden enlazarse geográficamente. Este concepto simple pero altamente poderoso y versátil ha probado ser crítico en la resolución de muchos problemas, que van desde el rastreo de vehículos de reparto, registrando los detalles de la aplicación de planificación, hasta la modelización de la circulación atmosférica global".*

54. David Comas y Ernest Ruiz, *Fundamentos de los Sistemas de Información Geográfica.* Barcelona: Ariel, 1993.

55. Stan Aronoff, *Geographic information systems: a management perspective.* Ottowa: WDL Publications, 1989.

56. FAO *(Organización de las Naciones Unidas para la Agricultura y la Alimentación)*: http://www.rlc.fao.org/es/prioridades/transfron/sig/intro/compo.htm [accedido 7 Julio 2011].

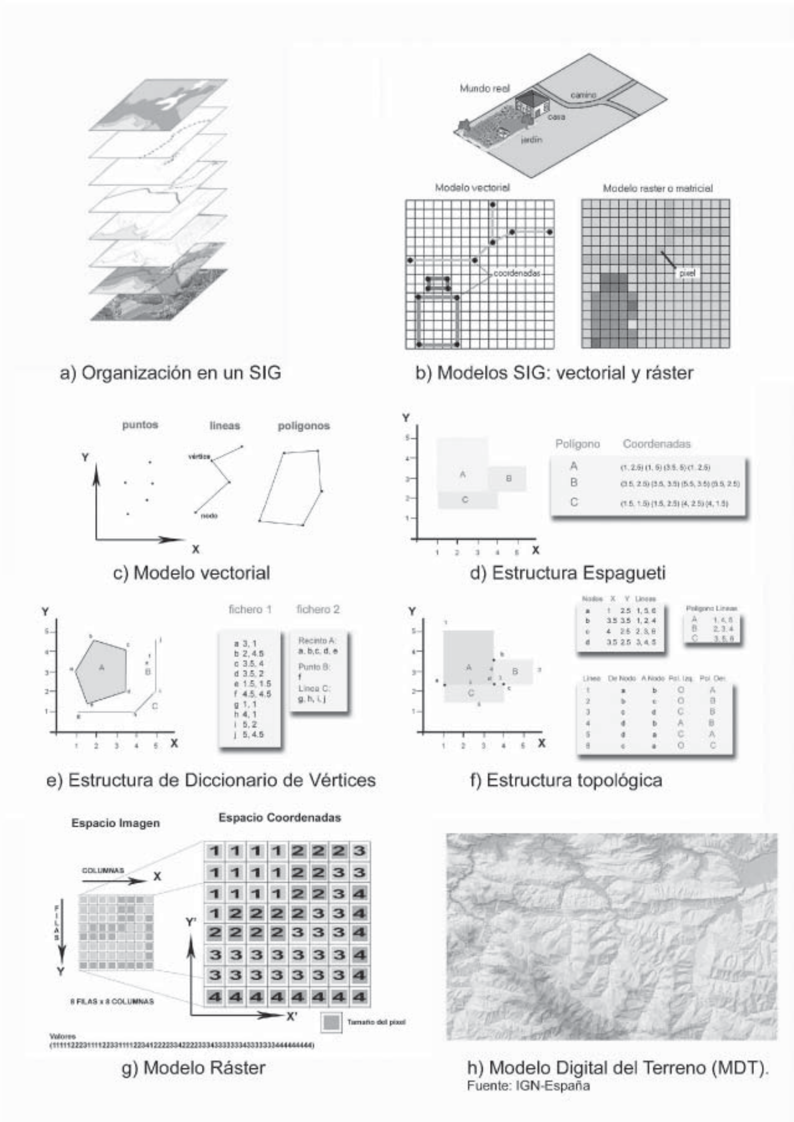


Figura 2. Aspectos clave de los SIG.

3. Modelos de Datos

A pesar de la heterogeneidad de la información, existen dos aproximaciones básicas para simplificar y modelizar el espacio, de modo que éste pueda ser almacenado y manipulado en un sistema informático, dando lugar, por tanto, a dos modelos de datos resultantes como se aprecia en la Figura 2b: el *modelo vectorial*, que representa la realidad en forma de puntos, líneas y polígonos, y es habitualmente utilizado para tratar fenómenos geográficos discretos (vías de comunicación, tejidos urbanos, cubiertas vegetales, etc.) y el *modelo ráster* o matricial, que representa los elementos geográficos mediante celdas regulares (cuadrados o rectángulos), usado generalmente para representar fenómenos continuos. Ambos sistemas son complementarios y conviven dentro de los SIG, aunque cada uno de ellos resulte más o menos apropiado para el estudio de un tipo de información específica.

3.1. El Modelo Vectorial

En el sistema vectorial, para la descripción de los objetos geográficos se utilizan vectores, utilizando una notación explícita por medio de coordenadas que se corresponden con el/los punto/s que definen espacialmente cada una de las entidades, a las cuales se asocian sus atributos temáticos (ver Figura 2c). Para ello se utilizan las denominadas primitivas geométricas de dibujo, referenciadas a un sistema de coordenadas geográfico conocido.

En este tipo de modelo SIG, un punto se codifica en la base de datos por medio de un par de coordenadas (x, y) que definen, inequívocamente, su posición. Una línea, es una secuencia de coordenadas (x, y), donde los puntos de principio y final de línea se denominan nodos y los puntos intermedios son los vértices. Los polígonos o superficies están representados por una agrupación de líneas cerradas, de tal manera que el primer punto es igual al último del elemento poligonal.

Dentro del modelo vectorial se pueden formular varias estructuras que son, en definitiva, distintas representaciones y formas de almacenar los datos dentro del mismo modelo. Las más importantes son: el *modo espagueti*, el *diccionario de vértices* y la *estructura topológica*[57].

57. Se puede consultar, por ejemplo: Tor Bernhardsen, *Geographic information systems: an introduction.* New York, NY: Wiley, 2002.

3.1.1. Estructura Espagueti

Es la estructura más simple de los modelos vectoriales y consiste en almacenar los datos como colección de puntos y segmentos lineales sin establecer relaciones lógicas entre ellos. Registra la geometría de los objetos, pero existen duplicidades de vértices y los límites comunes de polígonos adyacentes aparecen registrados varias veces, las líneas se superponen pero no se cruzan, y no existen puntos de intersección entre ellas (Figura 2d).

El modo espagueti, o de listado de coordenadas como también se le conoce, presenta claras desventajas al carecer de las interrelaciones entre entidades geográficas, al no ser una estructura topológica y debido a la redundancia de datos hace que no sea posible la realización de análisis espacial y es, igualmente inadecuado, para la visualización y representación cartográfica.

3.1.2. Diccionario de Vértices

Este tipo de estructura elimina los problemas de redundancia del modo anterior, ya que registra una sola vez las coordenadas de cada vértice posibilitando la ubicación espacial de manera inequívoca, al mismo tiempo que crea un diccionario de los puntos que constituyen cada objeto espacial almacenado en la base de datos geográfica (ver Figura 2e). Mejora la organización de los datos pero no la topología, que sigue sin estar suficientemente definida.

3.1.3. Estructura Topológica

Más sofisticados son los modelos topológicos vectoriales, donde las relaciones y conexiones entre los objetos son almacenadas, de forma conjunta, en la base de datos. Esta estructura es la más adecuada, en terminología SIG, para la realización de manipulaciones, análisis espacial y representaciones cartográficas posteriores.

El modelo topológico está basado en la teoría matemática de grafos, y emplea *nodos* y *aristas o arcos* como elementos de conexión

(ver Figura 2f). Por ejemplo, dos líneas que se cruzan se intersectan obligatoriamente formando un nodo; cada arco comienza y termina en un nodo, indicando cuál es el de inicio y el de fin de arco, independientemente de los vértices que puedan definir geométricamente cada línea; un polígono está definido por varios arcos con sus nodos respectivos, sin que exista duplicidad de bordes con sus vecinos y donde se almacena, en base de datos, cuáles son los polígonos que se encuentran a su derecha e izquierda.

Las ventajas de la estructura topológica se hacen evidentes pensando en las posibles consultas complejas que podrían realizarse a la base de datos, sólo este modelo soporta funciones de superposición geométrica, análisis de redes, determinación de caminos óptimos o generación de áreas de influencia, entre otras.

3.2. El Modelo Ráster

El modelo ráster divide el espacio en una serie de celdas regulares (cuadrados generalmente) en forma de filas y columnas. Este modelo codifica el interior de los objetos geográficos, registrando de forma implícita la frontera de los mismos, ver Figura 2b.

Cada celda de esta matriz o malla se denomina, también, píxel[58]. La resolución geométrica del modelo SIG va a depender del tamaño de estas celdas, cuanto más pequeño sea el píxel más precisa va a ser la definición digital de los elementos geográficos que representan. Por otro lado, el producto del número de filas por el de columnas define el tamaño de la matriz y, por tanto, el tamaño del fichero.

La topología está implícita en la regularidad de la rejilla y el establecimiento del origen de coordenadas (normalmente la esquina superior izquierda), de esta manera se conoce cuáles son los vecinos de una celda o píxel de la malla.

58. El término *"píxel"* proviene de la contracción de dos palabras en inglés *"picture element"* y define los elementos discretos, más pequeños, en los que se divide una imagen digital.

Los datos ráster tienen una forma especial de definir la ubicación geográfica. Una vez que las células o píxeles han sido georreferenciados con precisión (con respecto a un sistema de coordenadas terrestre), es fácil obtener un listado ordenado de todos los valores de cada celda del ráster en una geodatabase o base de datos geográfica (ver Figura 2g).

Cada celda en un ráster, además de su posición, puede llevar información y atributos asociados muy diversos:

- *Información temática*: como por ejemplo datos de usos del suelo, información topográfica de elevaciones, imágenes de temperatura, etc.
- *Información espectral*: como imágenes de satélite y fotos aéreas. Cuando se trata de datos multibanda, cada píxel con una única posición, posee más de un valor asociado con él. Con múltiples bandas, cada una, por lo general, representa un segmento del espectro electromagnético recogido por un sensor, incluidas las series que no son visibles al ojo humano, como las secciones de infrarrojos o microondas.
- *Features*: puntos, líneas y polígonos, en formato imagen, como mapas escaneados y fotografías.

Los modelos ráster en un SIG generan archivos muy voluminosos, por lo que todos los sistemas poseen técnicas de compresión de datos para facilitar el almacenamiento de los mismos. El más simple de ellos es la codificación *"RunLength"*, donde el sistema almacena pares de números: el valor del dato y el número de veces que ese valor se repite. Esto puede reducir el tamaño de los archivos de manera significativa.

3.3. Modelos Avanzados: Modelos Digitales del Terreno.

Los modelos de datos expuestos hasta ahora, se limitan a describir el mundo real mediante el uso de entidades bidimensionales, pero existen otros modelos usados en SIG que pueden representar la superficie del terreno en tres dimensiones, el factor tiempo o aplicados a objetos móviles, por ejemplo[59].

59. Véase Tor Bernhardsen, *op. cit.,* p. (85).

La representación digital de la superficie del terreno, de enorme interés dentro de los SIG y de la Cartografía, es conocida como *modelo digital del terreno* (MDT) y/o *modelo digital de elevaciones* (MDE). Los MDT y MDE no son excluyentes de las estructuras básicas de modelos vectorial y ráster, sino que participan de ambos tipos, siendo más bien una extensión particular de los mismos.

Una de las definiciones más utilizadas es la que hace A. Felicísimo[60], este autor define un MDT como "*una estructura numérica de datos que representa la distribución espacial de una variable cuantitativa y continua".* Esta definición permite que la característica a representar no tenga por qué ser la altitud, dejando el término homólogo MDE para cuando se pretende trabajar, específicamente, con las alturas, estos últimos *"describen la distribución espacial de la altimetría del terreno mediante un conjunto de datos acotados".*

Las formas elementales de estructuración posible de este tipo de modelos avanzados son cuatro fundamentalmente, dos de ellos son vectoriales: isopletas o contornos y redes irregulares de triángulos (*TIN, triangulated irregular network*); los otros dos son ráster: malla regular y matrices jerárquicas escalables (en estructuras de tipo *quadtree*). Las más utilizadas, en los SIG, son los de tipo TIN y los de malla regular.

3.3.1. Redes de Triángulos Irregulares (TIN)

Las redes TIN consisten en un conjunto de triángulos irregulares adyacentes que se ajustan a la superficie del terreno. Éstos se construyen a partir de puntos de coordenadas *(x, y, z)* conocidas, que constituyen los vértices de los mismos. Se trata de una estructura en la que el terreno queda representado, en forma de mosaico, por el conjunto de las superficies planas que se ajusta mejor en cada uno de los triángulos definidos.

A cada triángulo se le puede asociar el valor de la pendiente y la orientación, que definirán la dirección del plano del triángulo.

60. Angel M. Felicísimo, *Modelos digitales del terreno. Introducción y aplicaciones en las ciencias ambientales.* Oviedo: Pentalfa, 1994.

Una característica de los TIN es que el tamaño de los triángulos se puede ajustar a la complejidad del terreno, de modo que éstos sean más pequeños en zonas con mayor variación de relieve, y más grandes en zonas con menor variación. Esto hace que el número de datos que se deben almacenar sea más reducido[61].

3.3.2. Malla Regular

Denominado también de matrices regulares (*URG, uniform regular grids*), esta estructura es el resultado de superponer una malla sobre el terreno y extraer la altitud media de cada una de las celdas de la matriz. La cuadrícula puede adoptar formas variadas pero la más utilizada es la de malla cuadrada regular, con filas y columnas equiespaciadas.

Las posibilidades de análisis de los MDT y MDE, mediante técnicas SIG, son numerosas. La utilización de los *atributos topográficos primarios* (los derivados directamente a partir de los datos de altura representados en el modelo) permiten obtener información de pendiente, orientación, curvatura del terreno, etc.; el uso de *atributos topográficos secundarios o compuestos* (aquéllos que implican combinaciones de los atributos topográficos primarios) posibilitan un conjunto extenso de aplicaciones, como por ejemplo las de caracterización física del territorio o la variabilidad espacial de algunos procesos superficiales: índices de erosión, pérdida de biodiversidad, desertificación, índices de humedad, etc.

4, Georreferenciación: Sistemas de Coordenadas y Proyecciones Cartográficas

4.1. generalidades

Existen diversas ciencias como son: la Astronomía, la Geodesia, la Geofísica, la Topografía y la Cartografía, que se ocupan del estudio, medición y representación de la posición, la forma y las dimensiones

61. José Antonio Martínez Casasnovas, 'Modelos digitales de terreno: Estructuras de datos y aplicaciones en el análisis de formas del terreno y en Edafología.', *QUADERNS DMACS*, 25 (1999) pp. (51).

de la Tierra. Estas disciplinas aportan el conocimiento de la realidad física del territorio, necesario para la ubicación de los datos, fenómenos e información geográficamente referenciada.

Cualquier lugar del cielo o de la Tierra está determinado por unas coordenadas únicas respecto de un sistema de referencia que le distingue de los demás; la dificultad estriba en la representación de esos puntos en una superficie plana como es un mapa, asumiendo las consiguientes distorsiones o deformaciones debidas a la curvatura terrestre[62]. No obstante, se debe intentar preservar, en la medida de lo posible, las propiedades métricas del territorio, de tal forma que el mapa responda adecuadamente para el propósito que se persigue, y éste se realice de una forma precisa y rigurosa.

Si se desea representar una pequeña porción del terreno donde la escala va a ser muy grande, se puede considerar sin incurrir en errores significativos, la Tierra como plana, y por tanto se podrá utilizar un sistema cartesiano local de coordenadas, para las ubicaciones geográficas. Por el contrario, cuando se pretende representar una porción del territorio suficientemente extensa, como una región, un país, un continente o todo el planeta, donde la escala va a ser necesariamente muy pequeña, la Tierra ya no puede considerarse plana. En este caso, habrá que recurrir a artilugios matemáticos para representar la realidad del terreno, tridimensional y necesariamente afectada de la curvatura terrestre, en una superficie bidimensional, como es un mapa; esto se consigue mediante las denominadas *proyecciones cartográficas***.**

Las proyecciones cartográficas son, por tanto, funciones matemáticas de transformación biunívoca de las coordenadas geodésicas (también denominadas geográficas), de latitud y longitud obtenidas directamente de la superficie terrestre, en coordenadas cartesianas planas (o de mapa, X e Y).

El método analítico en el que se basan este tipo de transformaciones, requiere el conocimiento y definición de una serie de parámetros que definan inequívocamente la forma y tamaño de la superficie terres-

62. Fernando Martín Asín, *Geodesia y Cartografía matemática.* Madrid, 1990.

tre. Debido a la imposibilidad de materializar la Tierra en una expresión matemática, se recurre a distintas superficies de aproximación. En una primera aproximación, se puede considerar la Tierra como una esfera, donde el radio de la esfera será el parámetro, conocido, que servirá como constante de entrada en el cálculo de la proyección cartográfica.

A lo largo de la historia, a medida que se fue avanzando en el conocimiento científico y desarrollando nuevos instrumentos y técnicas de medida, que permitieron definir mejor la forma y dimensiones terrestres, se comprobó que la Tierra estaba ligeramente achatada por los polos[63], y la figura geométrica, de mejor ajuste, era un elipsoide de revolución. En lugar del radio de la esfera, los parámetros necesarios para el cálculo de la proyección, serán, en este caso, los semiejes del elipsoide.

Pero la superficie física de la Tierra no es lisa ni homogénea, está repleta de llanuras, montañas, océanos, valles, etc., esto es lo que se conoce como "superficie topográfica". Para hacer posible el cálculo y medición del relieve es preciso definir una superficie altimétrica de referencia. Esta es el ***geoide:*** esferoide tridimensional que resulta de imaginar una superficie de equilibrio representada por los océanos en calma, sin considerar el relieve terrestre, supuestamente prolongados por debajo de los continentes. Geométricamente, el geoide es el modelo que mejor determina la forma y la figura de la Tierra y físicamente se define como la superficie equipotencial del campo gravitatorio terrestre[64].

La expresión matemática del geoide es muy compleja para ser utilizada en cartografía, por lo que sólo se considera a efectos de determinación de altitudes[65]. Para simplificar el problema, como referencia planimétrica y para el cálculo de cualquier tipo de proyección cartográfica, que posibilite la elaboración de mapas y planos, y su extensión para los

63. *La forma de la Tierra. Medición del meridiano (1736-1744).* Madrid: Museo Naval, 1987.

64. Weikko A. Heiskanen y Helmut Moritz, *Geodesia Física.* Madrid: IGN, Instituto Geográfico Nacional, 1986.

65. Hay que reseñar que las determinaciones altimétricas clásicas en Geodesia y Topografía están referidas al geoide, estas se denominan "alturas ortométricas"; en tanto que los sistemas GPS refieren sus altitudes respecto del *elipsoide de referencia WGS-84* y se denominan "alturas elipsoidales".

sistemas cartográficos digitales como los SIG, se asimila la Tierra a una superficie regular, matemáticamente definida, como es el elipsoide de revolución**,** que es una superficie próxima al geoide y con una ecuación más sencilla; esto es lo que se denomina: *elipsoide de referencia.*

Dada la dificultad que supone la determinación de un único elipsoide de referencia a nivel mundial, los distintos países fueron adoptando aquel que mejor se asimile al geoide para su región geográfica concreta, sufriendo, además, sucesivas modificaciones en la determinación de sus parámetros a lo largo del tiempo; de forma que las medidas efectuadas por cada país se refieren a elipsoides de referencia distintos, dificultando enormemente las conexión de los trabajos a nivel internacional.

La determinación de un marco de referencia geodésico para cada país, obliga a determinar con exactitud los parámetros del elipsoide de referencia (semieje mayor, semieje menor, achatamiento y excentricidad) y a definir un punto, denominado *punto astronómico fundamental* sobre la superficie de la Tierra, en el que ambos modelos terrestres (geoide y elipsoide) son tangentes (coinciden la vertical del geoide y la del elipsoide en este punto y por tanto las coordenadas astronómicas y geodésicas respectivamente). Las coordenadas del punto fundamental, también denominado "punto origen" se determinan por métodos astronómicos, estableciendo de esta manera, unos datos iniciales a partir de los cuales se derivarán el resto. Elipsoide de referencia y punto fundamental constituyen lo que se denomina *Datum* y este es, en definitiva, el *sistema de referencia* para el cálculo y determinación de coordenadas y la georreferenciación de todos los puntos y datos geográficos.

La tecnología por satélite ha hecho posible la obtención de un Datum ubicado en el centro de masas, donde el centro del elipsoide de referencia y el punto origen coinciden con el centro de la Tierra; de este modo, elipsoide y geoide no están muy separados respecto a cualquier punto de la superficie y, aunque no "casan" perfectamente, las separaciones entre ambos se hayan uniformemente distribuidas.

En la Cartografía y la Geodesia española, el sistema de referencia oficial, utilizado hasta la actualidad, ha sido el que se conoce como

ED50 (*European Datum 1950*) que adopta el elipsoide de referencia de *Hayford 1910* (o Internacional) y como punto fundamental la Torre Helmert en la ciudad alemana de *Postdam.* A partir del año 2007, y por real decreto, se produce el cambio del Sistema Geodésico de Referencia en España, se establece como oficial el denominado ETRS-89 (*European Terrestrial Reference System 1989*) que utiliza un elipsoide de referencia geocéntrico, el denominado *GRS80*[66]. El plazo para que todos los organismos adapten su cartografía a este nuevo sistema es el 1 de enero del 2015. Hasta esta fecha, los organismos pueden realizar y publicar cartografía en cualquiera de los dos sistemas de referencia (ED50 ó ETRS89).

4.2. Coordenadas geográficas

La manera de posicionar un punto sobre la superficie de la Tierra, se hace mediante dos valores de medidas angulares, longitud (**λ**) y latitud (**φ**), valores referidos en un sistema geocéntrico de coordenadas esféricas.

La longitud, representada por el símbolo griego **λ,** es el ángulo diedro, formado desde el centro de la Tierra sobre el plano del ecuador, entre el meridiano que contiene el punto y el meridiano tomado como origen; ver Figura 3. El plano del Ecuador es el que pasa por el centro de la Tierra y es perpendicular al eje de rotación. Los meridianos son las líneas formadas por todos los puntos de igual longitud, y representan la intersección de la superficie de la Tierra con planos perpendiculares al plano del ecuador y que contienen el eje de rotación. Se ha definido un meridiano internacional de referencia o central, como origen de medida angular para las longitudes, que es el que pasa por el observatorio astronómico de Greenwich en Reino Unido. Por convención, la medida de longitud se expresa en grados sexagesimales y oscila entre 0º y 180º, resultando positiva y creciente hacia el oeste del meridiano de Greenwich y negativa hacia el este del meridiano origen.

66. A nivel práctico, el GRS80 puede asimilarse con el WGS84, que es el elipsoide de referencia utilizado en el Datum de EEUU, e implementado en los sistemas GPS y la cartografía global en Internet, como *Google Earth* o *Google Maps*. Los parámetros matemáticos de ambos elipsoides son muy similares, los semiejes mayores son idénticos y los semiejes menores difieren en décimas de milímetro.

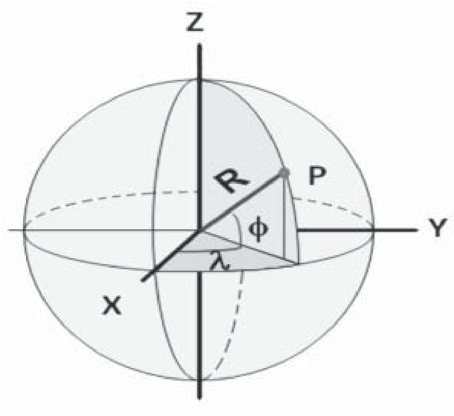


Figura 3. Sistema de Coordenadas Geográficas.

La latitud, representada por el símbolo **φ,** es el ángulo formado, desde el centro de la Tierra sobre el plano de meridiano, por la normal al elipsoide en el punto considerado y el plano del Ecuador, ver figura 3. Todos los puntos de la superficie de la Tierra con igual latitud definen las líneas de paralelos. La medida se expresa en notación sexagesimal partiendo del Ecuador, positiva hacia el norte (0º a 90º) y negativa en el sur (0º a -90º)[67].

4.3. Tipos de proyecciones cartográficas

Los sistemas geográficos de coordenadas no resultan demasiado cómodos a la hora de trabajar ya que utilizan medidas angulares de trigonometría esférica. Por ello, se adoptan sistemas de coordenadas rectangulares planos, más sencillos a efectos prácticos. Sin embargo, el problema añadido es que el elipsoide de revolución utilizado como modelo de referencia no es una superficie desarrollable, o sea, que no puede extenderse en un plano sin sufrir distorsiones. Esto hace, que sea

67. La notación angular de las coordenadas geográficas puede darse como grados y fracción decimal de grado (ejemplo: 43,9834ºN) o como grados, minutos y segundos (ejemplo: 18º 23’ 56’’N).

preciso recurrir a funciones de transformación de coordenadas (proyecciones cartográficas):

$$X = f(\lambda,\varphi) \qquad \lambda = F(X,Y)$$
$$Y = g(\lambda,\varphi) \qquad \varphi = G(X,Y)$$

La formulación matemática de las funciones (f, g, F, G) define las propiedades geométricas de la representación elegida y la correspondencia entre puntos homólogos en el mapa y en el terreno[68]. Para ello, se utiliza la red de paralelos y meridianos como malla auxiliar.

Se suelen clasificar las proyecciones en función de la geometría de la superficie sobre la que se realiza la proyección (planos, cilindros o conos), así como la disposición entre dicha superficie de proyección y la superficie terrestre (ver Figura 4), pudiendo ser tangentes o secantes:

- *Azimutales:* cada punto de la superficie de la Tierra se proyecta sobre un plano tangente en un punto al globo. Si este punto es el polo terrestre, se dice que son "*proyecciones azimutales polares*" y si está situado en el ecuador son "*ecuatoriales*". Dependiendo del origen del haz proyectivo, desde donde se realiza la proyección, se pueden diferenciar proyecciones: *Ortográficas* con origen en el infinito, haces paralelos; *Estereográficas* con origen del haz en el punto de la superficie de la Tierra opuesto al de tangencia del plano; y *Gnomónicas* con origen del haz en el centro de la Tierra. Estas proyecciones no permiten representar la totalidad del planeta y se recurre a dividir la Tierra en los dos hemisferios. Una de las proyecciones azimutales más utilizadas es la "Estereográfica Polar", sobre todo para las regiones geográficas polares.
- *Cónicas:* estas utilizan como plano proyectado, el desarrollo de un cono, normalmente con vértice en la prolongación del eje de la Tierra. En este caso, los meridianos son líneas equidistantes y convergentes y los paralelos arcos de círculo con centro en el punto de convergencia de los meridianos. La más utilizada y conocida es la "Proyección cónica conforme de Lambert", muy utilizada

68. El elipsoide es la superficie de referencia utilizada para los cálculos, y sólo se adoptará como superficie de referencia la esfera cuando ésta sea la más aconsejable para el caso de estudio concreto.

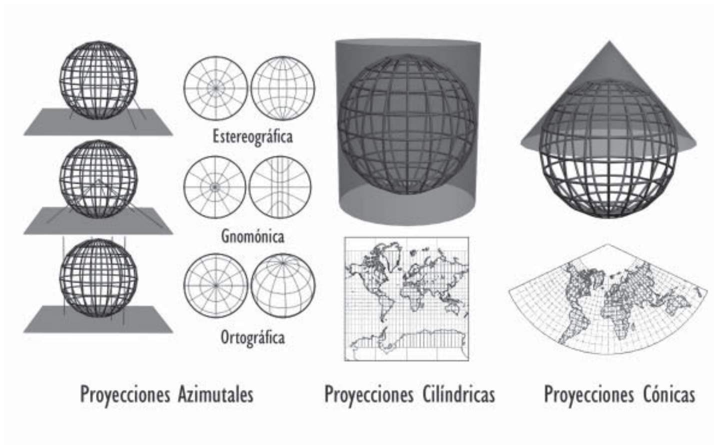


Figura 4. Tipos de Proyección. Clasificación básica.
Fuente: Adaptación de http://es.wikipedia.org/wiki/ProyecciónCartográfica

para la representación completa de países situados en zonas intermedias entre los polos y el ecuador.

- *Cilíndricas:* las más utilizadas son las ecuatoriales, en las que el plano proyectado es el desarrollo de un cilindro tangente a la superficie de la Tierra en la línea del ecuador. La característica principal de estas proyecciones es que los paralelos y meridianos son líneas rectas y se cortan entre sí en ángulos rectos. La proyección cilíndrica más utilizada a lo largo de la historia ha sido la *Proyección Mercator*; por ser conforme y mantener los ángulos, ha sido la elegida para la cartografía náutica, ya que permite el trazado directo de rumbos.

Existen otras proyecciones muy utilizadas, sobre todo para representar la totalidad de la Tierra completa, que aunque tienen como base la clasificación anterior, se han modificado para compensar de alguna forma las distorsiones propias de cada proyección a costa de forzar curvas o romper la continuidad del mapa en zonas de océanos para preservar los continentes. Como ejemplo, podemos destacar las proyecciones de: Bonne, Mollweide, Goode, Sinusoidal, Robinson, Van der Grinten, Eckert, etc. La *National Geografhic Society* ha adoptado la Proyección Winkel-Tripel como estándar para la representación de mapasmundi.

Es imposible representar un elipsoide sobre un plano sin distorsiones. Por ello, existe una gran diversidad de proyecciones que intentan transferir la información del globo al plano con deformaciones mínimas. Dependiendo del tipo de deformación se pueden clasificar en:

- *Conformes:* conservan los ángulos entre meridianos y paralelos, mantienen las formas.
- *Equivalentes:* mantienen constantes las relaciones de superficie.
- *Equidistantes:* mantienen las relaciones de distancia.

Tampoco es posible mantener las tres propiedades anteriores a la vez, sólo se puede conservar una de ellas o establecer un compromiso entre las tres, esto dependerá de la utilidad a la que sea destinado el mapa.

4.3.1 La proyección UTM

Esta proyección merece una mención especial por tratarse de una de las más utilizadas en la mayoría de los países y por ser la proyección oficial en España.

UTM son las siglas del inglés *Universal Trasverse Mercator*. Esta proyección se construye proyectando todos los puntos en un cilindro transversal, tangente al globo terráqueo a lo largo de un meridiano que se elige como meridiano de origen. A medida que se aleja del meridiano origen, las deformaciones aumentan, por lo que se recurre a subdividir la superficie terrestre en 60 zonas iguales, denominada *Husos*, de seis grados de longitud cada uno, con la cual resultan 60 proyecciones UTM para todo el planeta, cada una con su propio meridiano central de tangencia.

Las deformaciones son mínimas en latitudes cercanas al ecuador, pero a medida que se aproxima a latitudes polares, estas aumentan. Por ello, se limita su utilización a las zonas terrestres comprendidas entre los 84º de latitud norte y los 80º de latitud sur.

Para minimizar las deformaciones lineales que se producen al separarse del meridiano central y del ecuador, se recurre a aplicar un fac-

tor de reducción, de modo que la posición del cilindro de proyección en lugar de tangente se haga secante, generando dos líneas paralelas, a lo largo de las cuales, la deformación lineal o anamorfosis es la unidad.

Las ventajas de esta proyección son evidentes: es conforme (conserva los ángulos) y no distorsiona las superficies en grandes magnitudes; por ello es universalmente utilizada, fundamentalmente para usos militares.

En países, como el caso de España, que ocupa los Husos 29, 30 y 31(en península y Baleares) y el huso 28 para Canarias (ver Figura 5), esta proyección plantea algunas desventajas, como la de tener coordenadas duplicadas en los puntos de límite de Husos; no poder representar en esta proyección todo el país; la elección del huso o la duplicidad de las hojas cartográficas limítrofes entre husos. Estas desventajas han sido paliadas, en parte, por las aplicaciones SIG, que realizan complejos cálculos de transformación, y permiten, en tiempo real, visualizar gran cantidad de información gráfica independientemente de la proyección y del sistema de referencia en el que estén almacenados los datos.

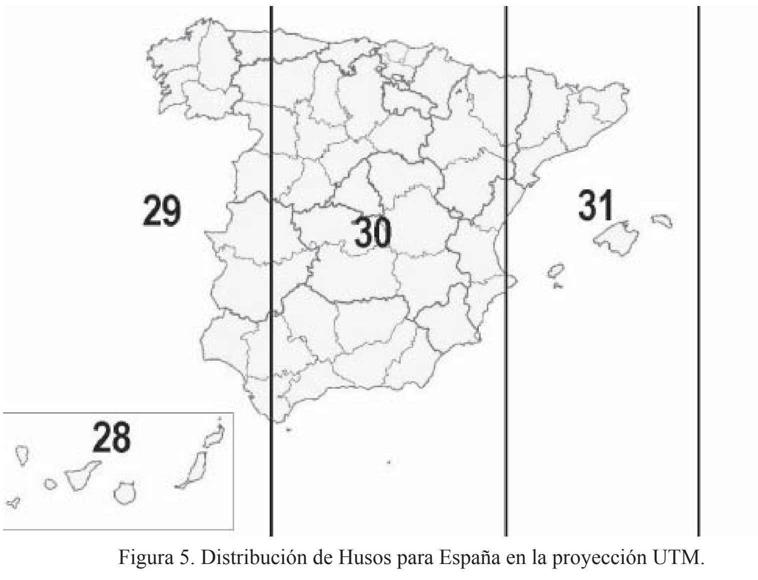


Figura 5. Distribución de Husos para España en la proyección UTM.

Capítulo II
Diseño y modelado de datos

Este capítulo describe los primeros pasos que hay que dar a la hora de afrontar el desarrollo de un Sistema de Información (SI) en general, o de un SIG en el caso más específico que nos ocupa, en concreto las fases que se refieren al diseño y modelado del conjunto de datos.

El Ciclo de Vida de un Proyecto

En líneas generales, el ciclo de vida de desarrollo de cualquier SI, independientemente de que se integren o no *datos geográficos* en él (en caso afirmativo nos referiríamos a un SIG), se compone de una serie de pasos secuenciales que interactúan entre sí. La división en fases que hace MÉTRICA v.3.[69] es la siguiente: el estudio de viabilidad del SI (análisis de requerimientos, problemas y restricciones, objetivos,…), su análisis, el diseño (incluye modelo conceptual, lógico y físico), su construcción (selección de hardware y software, creación de la estructura de la base de datos, conversiones de datos) y, por último, la implantación y aceptación del mismo (sistema en producción, formación de los usuarios, actualizaciones de producto, copias de seguridad,…).

Independientemente de la metodología elegida, todo este proceso se facilita mediante el empleo de las denominadas herramientas CASE (*computer-aided software engineering*), aplicaciones informáticas que

69. MÉTRICA v.3 es una metodología desarrollada por el Ministerio de la Presidencia de España para sistematizar las actividades que dan soporte al ciclo de vida del software. Está enmarcada dentro de la norma ISO 12207 relativa a *"Information technology – Software life cycle processes"*.

reducen los costes de tiempo y económicos en el desarrollo de un proyecto al facilitar la generación de toda la documentación del ciclo de vida del mismo.

A continuación, nos centraremos básicamente en las dos primeras fases, la del estudio inicial o de viabilidad del sistema y la de diseño, mientras que la creación de la estructura de la base de datos y la posterior carga de datos se tratarán con profundidad en el capítulo siguiente.

1. Fase de Viabilidad del Sistema

Es la primera fase en la implementación de cualquier sistema de base de datos. La conceptualización de la problemática a estudiar pasa necesariamente por abordar un proceso iterativo mediante el que se han de definir los *requerimientos de los usuarios* de forma progresiva, traduciendo éstos posteriormente a especificaciones técnicas para poder ser implementadas. No podemos olvidar que el sistema que diseñemos ha de satisfacer las necesidades de los usuarios finales del mismo. El objetivo último ha de ser definir unas especificaciones de producto lo más estables y completas posibles, aunque flexibles a los cambios derivados de la evolución en el tiempo de los sistemas y de la propia tecnología.

A la hora de definir los requisitos reales de un sistema es fundamental la participación de todas las personas involucradas en el proyecto, tanto la de los citados usuarios finales como la del personal técnico encargado del desarrollo del mismo (analistas, desarrolladores, etc.).

Los requerimientos de un sistema se suelen modelar mediante los denominados *casos de uso* y muestran, mediante una notación gráfica, los pasos que debería dar un determinado *actor*[70] para realizar una tarea concreta, es decir, la forma en que interactuaría éste con el sistema, y cuál sería la respuesta obtenida.

70. Un *actor* es una entidad externa al sistema que guarda una relación con éste y que le demanda una funcionalidad concreta.

A modo de ejemplo, el diagrama de casos de uso de un SIG genérico podría asemejarse al de la Figura 6. En este caso los *actores*, representados como muñecos fuera del cuadro que delimita el sistema, hacen referencia a los distintos roles de los usuarios, diferenciando entre usuarios expertos y no expertos, y otro usuario genérico proveedor de los datos. Los *casos de uso* se representan como elipses y corresponden a distintas funcionalidades de la aplicación: búsquedas básicas y/o avanzadas (según un ámbito geográfico, un atributo temático o un intervalo de tiempo; en realidad las tres componentes de los datos geoespaciales), generación de informes, visualización de los datos, etc. Por último, los casos de uso pueden interactuar entre sí mediante diferentes tipos de *relación* representados por líneas continuas o discontinuas (generalización, inclusión y/o extensión).

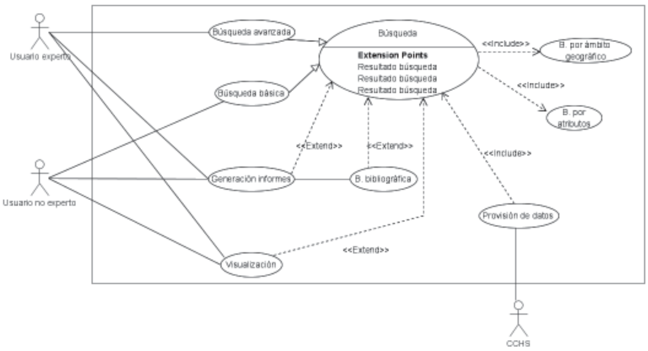


Figura 6. Diagrama de casos de uso de un SIG de propósito general.

2. Fase de Diseño

El *modelado* es una fase fundamental e imprescindible en la creación de cualquier sistema de calidad que implique la implementación de una base de datos sea cual sea su naturaleza, y el modelo en sí es su materialización, la forma en que se organizarán los datos. También es el instrumento que ha de servir, ante una problemática concreta, como medio de comunicación entre todas las personas implicadas, usuarios y

desarrolladores (es muy importante recoger esta pluralidad de perspectivas); mediante él se controla en definitiva la estructura y arquitectura de un sistema o su funcionamiento, entre otros.

Un *modelo* es una abstracción y simplificación de la realidad y existen diferentes lenguajes formales para expresarla, siendo uno de los más aceptados el denominado *lenguaje unificado de modelado (UML)*[71] que utiliza una serie de notaciones gráficas y diferentes tipos de diagramas en función de la parte que se esté tratando del sistema: la parte estructural, la de comportamiento o la arquitectural[72]. El modelo es tan solo una descripción conceptual de la base de datos y no hace referencia en ningún caso a la estructura física de la misma, que se abordará en un paso posterior en la definición de lo que se conoce como *esquema* de la base de datos, que ya sí define las tablas y sus atributos, y también las relaciones entre ellas.

La fase de diseño pasa por varios tipos de modelado que se reducen básicamente a tres: conceptual, lógico y físico. A continuación, resumiremos brevemente cada uno de ellos.

2.1. Modelo de datos conceptual

El modelo de datos *conceptual* representa la realidad a altos niveles de abstracción, y se apoya en la definición de los requerimientos de usuario que se especificaron en la fase anterior correspondiente al estudio de viabilidad del sistema. Es independiente del hardware a utilizar y del sistema gestor de base de datos (SGBD), que es el software que almacena los datos, los manipula y los recupera tras una consulta a la base de datos[73], también del paradigma de modelado a implementar (relacional, orientado a objetos, u objeto-relacional). Su objetivo es definir contenidos, no estructuras de almacenamiento, y debe reflejar

71. Unified Modeling Language (UML), es un lenguaje gráfico para visualizar, especificar, construir y documentar un sistema.

72. Grady Booch, *El Lenguaje Unificado De Modelado /: Guía Del Usuario*, 2o edn. Addison Wesley, 2006.

73. David J. Maguire y Michael F. Goodchild, *Geographical Information Systems: Principles and Application*, ed por David W. Rhind, 1o edn. Essex: Longman Scientific and Technical, 1991.

todos los aspectos relevantes del mundo real a modelar; en este sentido, es el modelo más "cercano" al punto de vista del usuario.

Uno de los modelos conceptuales más extendidos, aplicado generalmente, aunque no siempre, a los modelos de tipo relacional, es el denominado *Entidad-Relación (ER)*[74], llamado así porque muestra los datos en términos de entidades y de sus relaciones, en contraposición con los modelos orientados a objetos (OO) que agrupan los datos conceptualmente en las denominadas "clases" que veremos posteriormente. En un modelo ER se entiende por:

- *entidad,* un objeto abstracto del cual se pueda almacenar información en una base de datos (es equivalente a una instancia en los modelos orientados a objetos, ej.: una determinada parcela del catastro de rústica, un yacimiento arqueológico concreto, un navío que realizaba rutas comerciales…),
- *conjunto de entidades,* una colección de objetos similares (equivalente al concepto de clase en OO, ej.: "parcela", "yacimiento" o "navío"),
- *atributos,* las propiedades de esas entidades (ej.: en una parcela serían la referencia catastral, el propietario, la superficie, el tipo de uso del suelo, etc., o en un yacimiento, su denominación, su perímetro, la época a la que pertenece, etc., y con respecto a un navío, su nombre, su tipología, o su tonelaje),
- *relaciones,* las conexiones entre dos o más conjuntos de entidades (ej.: una parcela puede pertenecer a varios propietarios, o un propietario puede ser dueño de varias parcelas),
- *dominio,* el conjunto de valores que puede tomar un determinado atributo (ej.: para una clase "uso del suelo" el dominio sería el listado de valores posibles: forestal, urbano, cultivo, lámina de agua, etc.).

En el modelo ER (ver Figura 7) los *conjuntos de entidades* se representan como rectángulos dentro de los cuales se escribe el nombre de la entidad (que suele ser un sustantivo), y las relaciones entre enti-

74. Peter Pin-Shan Chen, 'The Entity-Relationship Model: Towards a unified view of Data', *ACM Transactions on Database Systems (TODS)*, 1 (1976) pp. (9-36).

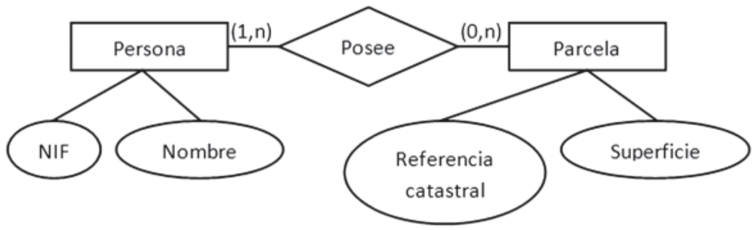


Figura 7. Modelo ER

dades se representan como rombos que las conectan en cuyo interior se escribe el nombre de la relación (que suele ser un verbo).

El concepto de relación va unido al de *cardinalidad*, entendiéndose por ésta el número de relaciones en las que una entidad puede aparecer; siempre se especifican tanto el valor mínimo como el máximo. La notación que se utiliza varía mucho pero una de las más extendidas es la que se muestra en el ejemplo anterior: una persona puede no ser propietario de ninguna parcela o poseer varias (relación 0,n), mientras que una parcela puede pertenecer a un propietario o a varios a la vez (relación 1,n).

Las entidades se caracterizan mediante sus *atributos*, que en el modelo ER se representan como elipses que se unen con líneas a las entidades y dentro de las cuales se escribe el nombre del atributo. Así, las personas quedan descritas a través de su NIF y su nombre, entre otros, mientras que las parcelas se definen con su referencia catastral o su superficie, todos ellos atributos de estas entidades. De los atributos posteriormente se especificará si son obligatorios u opcionales (en este caso admitirían valores nulos, *null*).

Para cada conjunto de entidades hay que definir lo que se denomina identificador o *clave*, un atributo cuyo valor es único para cada una de las entidades (en ocasiones puede estar formada por varios atributos a la vez), y generalmente se utiliza la notación id_<entidad> (en el ejemplo anterior para la persona utilizaríamos como clave el atributo

“id_persona”, que sería el NIF, o para la parcela el “id_parcela”, que sería la referencia catastral ya que ambos son valores que definen de forma unívoca cada entidad).

2.2. Modelo lógico

El modelo *lógico* parte del modelo conceptual y ya sí que depende del tipo de SGBD que se vaya a utilizar, aunque no del producto concreto. Define ya una estructura de la base de datos y es un modelo más depurado que el anterior.

Existen estándares de facto, como los que promueve el OMG (Object Management Group)[75] —una organización internacional que impulsa la utilización de estándares— en el diseño e implementación de sistemas como es el caso del ya mencionado UML[76], válido no solo para sistemas OO sino también para modelos relacionales. Ha tenido repercusiones importantes en el campo de los datos espaciales porque éstos son un tipo de datos complejo por sí mismo.

Como se ha adelantado anteriormente, existen básicamente tres aproximaciones al modelado en función del tipo de base de datos que vayamos a implementar en el sistema: el modelo relacional, el orientado a objetos y el objeto-relacional. A continuación veremos brevemente qué diferencia a unos de otros.

2.2.1. Modelo relacional

El concepto de *modelo de tipo relacional* fue introducido en los años 70 por Ted Cood y se basa en la creación de una serie de tablas donde cada fila es una entidad y cada columna un atributo que la caracteriza (de tipo numérico, textual, etc.).

El paso del modelo conceptual ER al modelo lógico de tipo relacional consiste en convertir las entidades a tablas, los atributos a

75. http://www.omg.org/
76. http://www.uml.org/

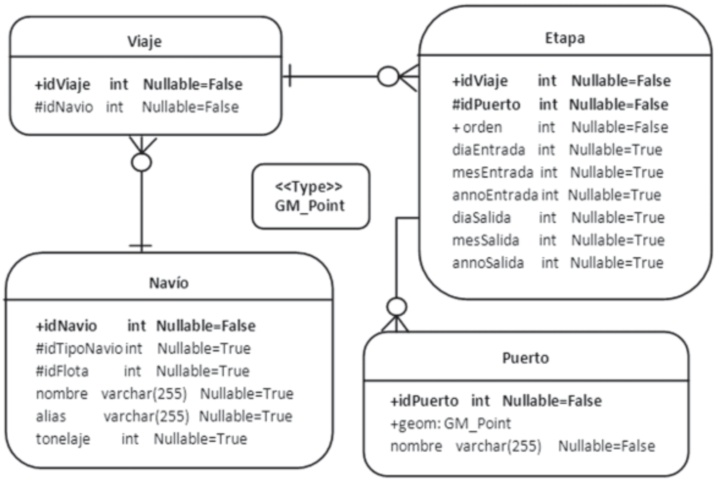


Figura 8. Ejemplo de modelo de datos.

columnas de las tablas, y los identificadores a claves primarias y claves foráneas. Hay que tratar los diferentes tipos de cardinalidades, por ejemplo, las del tipo varios a varios (n:m) se resolverán creando una tabla intermedia que contendrá las claves primarias de las dos entidades implicadas.

Para representar este tipo de modelos la mayoría de las herramientas CASE utilizan representaciones como la que se muestra anteriormente (ver Figura 8), en la que las tablas se representan como rectángulos con una cabecera donde se inserta el nombre de la tabla y una serie de filas con los nombres de los atributos; con las líneas que unen tablas se muestran tanto las relaciones como el tipo de cardinalidad (uno a muchos, uno a uno, muchos a muchos, etc.).

En este ejemplo se representan, en el caso de un SIG dedicado por ejemplo a representar y analizar las rutas comerciales marítimas de la Primera Edad Global[77], cuatro entidades relacionadas: los navíos, los

77. Período de la historia comprendido entre 1400 y 1800, objeto de estudio del proyecto DynCoopNet, liderado por Ana Crespo Solana del IH (CCHS-CSIC).

viajes que éstos emprendieron y las etapas de las que constaba cada uno de ellos, pasando por sus puertos. Cada entidad está representada por una caja donde se listan los atributos que las definen (con especificación en negrita de los campos clave, aquellos a través de los cuales las tablas se relacionan entre sí), y las relaciones de cardinalidad entre las entidades se representan con líneas (un navío puede o no haber realizado uno o varios viajes, y un viaje se puede componer de una o varias etapas).

Como última tarea en la definición del modelo lógico, hay que realizar lo que se conoce como *normalización*, un proceso cuyo objetivo es depurar el modelo optimizándolo para evitar por ejemplo la redundancia de datos. Un buen modelo debería satisfacer al menos lo que se denominan las tres formas normales de Boyce-Codd-FNBC[78], siendo cada una de ellas más restrictiva que la anterior. Así, la primera forma normal (1FN) implica que no haya atributos repetidos en las tablas, la segunda forma normal (2FN) implica la anterior y además que todos los atributos que no sean claves dependan únicamente de la clave primaria, y por último, la tercera forma normal (3FN), que obliga a que se cumplan las dos anteriores e introduce nuevas restricciones.

2.2.2. Modelo orientado a objetos

De la dificultad de modelar objetos o entidades complejas del mundo real (imágenes, datos espaciales,…) surgieron los sistemas basados en los llamados *modelos orientados a objetos*, mucho más abstractos que los relacionales. En un modelo OO:

- Un *objeto* es una abstracción, un concepto con límites definidos y con significado, a cada objeto individual se le conoce como *instancia* (equivale a una fila de una tabla en el modelo relacional).
- Las *clases* son colecciones de objetos similares con atributos compartidos (equivale a la entidad o tabla en el modelo relacional, ej.: una clase genérica poligonal puede ser el "uso de suelo", independientemente de que luego unos polígonos hagan referencia

78. Cristopher J. Date, *An Introduction to Database Systems.* Addison-Wesley, 2004.

a masas de agua, otro a zonas agrícolas, a zonas forestales, etc.). Las clases se organizan en una jerarquía donde unas son superclases y otras subclases.

- Los *atributos* son características que describen los objetos de una clase; cada atributo tiene un valor para cada instancia del objeto.
- Los *métodos* son las operaciones que pueden realizarse sobre un objeto (ej.: sobre una clase "yacimiento arqueológico" se podrían aplicar los métodos de borrar, dar de alta, editar, seleccionar, cambiar nombre, cambiar simbología,…).
- Entre clases se puede dar lo que se denomina *herencia*, de forma que las subclases "heredan" de las clases superiores tanto los atributos como el comportamiento.
- En un modelo OO se suelen implementar algunas propiedades como son la generalización, la asociación o la agregación de objetos, pues encaja mejor con la realidad a modelizar.

En el caso de los SIG tendremos por tanto *objetos geográficos* que pueden definirse como paquetes integrados formados por una geometría (tiene que ver con su posición en el espacio y su extensión), unas propiedades o atributos (que los describen) y unos métodos o acciones (que se pueden ejecutar sobre ellos). Objetos similares se agruparán en clases, existiendo por tanto clases de naturaleza geométrica y otras que no lo son. Un ejemplo de SGBD objeto - relacional es PostgreSQL.

Según lo visto, el modelo OO es más completo que el ER porque no solo define la parte descriptiva sino también el comportamiento. El modelo lógico también debería incluir un *diccionario de clases*, un documento que resumiera los elementos que constituyen cada clase: nombre de la clase, tipo, descripción y atributos (nombre, tipo de dato, valor inicial, su dominio en su caso, etc.).

2.2.3. Modelo objeto-relacional

Por último, los *modelos objeto-relacionales* tratan de ser un híbrido entre los dos anteriores, en realidad se podría decir que son una extensión del tradicional modelo relacional en un intento de solventar sus limitaciones, aunque no los desarrollaremos en detalle en este libro.

2.3. Modelo Físico

El modelo *físico* tiene ya que ver con la implementación en una base de datos; es el nivel de mayor concreción. Es el momento ya de especificar cosas como el SGBD concreto a utilizar que almacenará y manipulará los datos, la arquitectura del sistema (arquitectura distribuida y modelo cliente-servidor es lo más extendido), la localización física de los archivos de datos, el espacio que ocupará cada atributo dentro del campo (longitud y tipo), etc. Es por tanto un modelo ya más cercano al ordenador que a los usuarios, a diferencia de lo que ocurría con el modelo conceptual.

Una *base de datos*[79] es una colección de datos estructurados junto con el diccionario de datos que describe los datos, las reglas de integridad que evitan incongruencias, y los mecanismos de almacenamiento y consulta por parte de los usuarios.

En un SIG a través del modelo físico las clases del modelo lógico se implementarán a través de tablas en la base de datos, donde cada fila será una instancia del objeto y cada columna un atributo[80]. Por ejemplo, una clase genérica llamada "carreteras" podría incluir diferentes instancias como "autovía", "carretera nacional", o "carretera local".

Como ya hemos visto en el apartado anterior, los SGBD más habituales son bien los relacionales o bien los orientados a objetos, o incluso una combinación de ambos. No obstante, cada uno de ellos debería cumplir una serie de funcionalidades específicas[81] que se resumen a continuación:

- permitir tanto el almacenamiento y la recuperación de datos (mediante preguntas, *queries*) como la selección de determinados re-

---

79. Avi Silberschatz, Henry F. Korth y S. Sudarshan, *Database System Concepts*, McGraw-Hill series in computer science. Estados Unidos: Mcgraw-hill College, 1999.

80. United Nations Publication. Department of Economic and Social Affairs, *Handbook on Geospatial Infrastructure in support of the Census Activities*, FNew York: United Nations, 2009, ciii. <http://unstats.un.org/unsd/publication/seriesf/Seriesf_103e.pdf> [accedido 12 Mayo 2011].

81. Peter A. Burrough y Rachael A. McDonnell, *op. cit.*

gistros en base a una serie de campos (atributos) y relaciones,
- presentar una interfaz para la entrada y edición de los datos,
- implementar una serie de controles automáticos (reglas de integridad) que aseguren la consistencia de los datos, y
- disponer de mecanismos para evitar la manipulación y/o eliminación deliberada de datos.

En la mayoría, el acceso a los datos se realiza a través de un lenguaje de consulta determinado, siendo el conocido como SQL (Structured Query Language) el más extendido, o variantes de éste en el caso de los datos geográficos.

Independientemente del sistema que elijamos para nuestro desarrollo, es además esencial que el SGBD cumpla estándares con el fin de asegurar la interoperabilidad[82] entre sistemas; principio fundamental en el que se basan las denominadas Infraestructuras de Datos Espaciales (IDE), de gran desarrollo en los últimos años, cuyo objetivo es integrar a través de Internet los datos, metadatos, servicios e información de tipo geográfico que se producen a cualquier nivel (nacional, regional y local). Existe ya todo un desarrollo legislativo que regula este tipo de infraestructuras tecnológicas[83].

Algunas de las tareas a realizar en el modelo físico, relacionadas con la construcción del esquema de la base de datos son:

- Traducir los modelos conceptual y lógico a tablas, relaciones, e instancias (filas) de una clase concreta.
- Describir los tipos de datos de cada atributo (carácter, numérico, fecha, etc.), definir los que se convertirán en claves, decidir su obligatoriedad, es decir, si se permiten o no valores nulos, indicar los índices espaciales que se crearán para optimizar las búsquedas, especificar en su caso los dominios de los atributos, etc.

---

82. La interoperabilidad es la capacidad de los sistemas de información y de los procedimientos a los que éstos dan soporte, de compartir datos y posibilitar el intercambio de información y conocimiento entre ellos.

83. En el caso concreto de España existe la denominada IDEE (Infraestructura de Datos Espaciales de España). Consultar http://www.idee.es/.

- Describir las *constraints*, las reglas que gestionan un determinado tipo de relación; mecanismos de control de calidad de los datos automáticos.

Los Datos Geográficos

Para finalizar este capítulo, algunas reflexiones sobre las particularidades de los datos geográficos a implementar en un SIG, aunque se desarrollarán en profundidad en los capítulos siguientes.

Como ya se ha comentado, los datos geográficos o *geoespaciales* son complejos por su naturaleza, y la base de datos que los almacene ha de ser capaz de manejar dicha complejidad, ha de poder almacenar tanto *posiciones espaciales* (en base a un determinado sistema de referencia) como *geometrías*, pero también la *topología.*

Este tipo de datos requiere además una serie de operadores de consulta espaciales que en ocasiones no se resuelve con el lenguaje SQL estándar, existe por ejemplo una variante de éste en el caso de los modelos OO llamada OQL (Object Query Language), o para consultar datos espaciales se ha diseñado el llamado SSQL (Spatial SQL) —que introduce conceptos del algebra de ROSE (RObust Spatial Extension)[84]—, PSQL (Pictoral SQL) para los datos obtenidos por rasterización[85], u Oracle Spatial y PostGIS como extensiones para realizar consultas espaciales.

Tras el modelo conceptual, la materialización real de la información en una base de datos implica decidir, como ya se adelantó en la introducción, entre un *modelo de representación vectorial* (más adecuado para representar entidades geométricas discretas) o uno *ráster* (más adecuado para variables de tipo continuo). Cada uno de ellos maneja tanto la componente espacial como la temática de forma diferente. De

84. Ralf Hartmut Güting y Markus Schneider, 'Realm-based spatial data types: The ROSE algebra', *The VLDB Journal*, 4 (1995) pp. (243-286).
85. 'Base de datos espacial - Wikipedia, la enciclopedia libre', 2011 <http://es.wikipedia.org/wiki/Base_de_datos_espacial> [accedido 19 Mayo 2011].

cualquier forma, existen mecanismos para convertir los datos de un modelo de representación a otro, de forma que lo ideal es utilizar un SIG que sea capaz de manejar ambos formatos.

A diferencia de lo que ocurría en los SIG de años atrás, actualmente existe una tendencia a integrar en una misma base de datos los datos espaciales o cartográficos y los no espaciales, los atributos; en el modelo relacional estos dos grupos de datos estaban separados, incluso en bases de datos diferentes. Con el incremento de complejidad de los datos en un SIG, el modelo de datos relacional se ha quedado corto por no ser capaz de almacenar de forma óptima tanto los elementos geométricos y sus relaciones espaciales como las superficies, de ahí el cada vez mayor éxito de los modelos orientados a objetos, que posibilitan definir clases abstractas e incluir todos los datos en la misma base de datos (espaciales, topología, atributos e incluso metadatos).

Una vez decidido el modelo de representación de los datos (ráster o vectorial) solo queda decidir el tipo de almacenamiento digital de los mismos en el ordenador; como se adelantó en el capítulo de fundamentos existen varias alternativas para esto, aunque dependen directamente del software SIG a utilizar, siendo al final algo transparente para el usuario. Todos ellos tratan de minimizar el espacio de almacenamiento y maximizar la eficiencia en los cálculos.

Capítulo III
La Base de Datos Geográfica

En los capítulos anteriores hemos visto cómo abstraer y modelizar la información geográfica para poder utilizarla en un proyecto, pero esta información se tiene que introducir en un SIG, y es en este capítulo donde veremos cómo y de qué manera se puede almacenar esta información en bases de datos *especiales* —dada la peculiar naturaleza de la información geográfica—, y cuáles son las principales fuentes de datos geográficos, haciendo hincapié en las más utilizadas en los SIG en la actualidad.

Generación de una base de datos Geográfica (BDG)

"Una base de datos espacial es una colección de datos referenciados en el espacio que actúa como un modelo de la realidad" (NGCIA, 1990)[86].

La principal diferencia con una base de datos convencional se refiere a la naturaleza de la información que contiene. En una *base de datos geográfica (BDG)*, se almacenan entidades del mundo real con sus dos componentes según un modelo espacial, de forma estructurada y sin duplicidades. Estas entidades, abstractas o no, ocupan un lugar en el espacio según un sistema de referencia, mantienen unas relaciones espaciales entre ellas y con el resto de los elementos del entorno, y poseen unos atributos propios que las diferencian entre sí, en un tiempo

86. Jorge del Río San José, *Introducción al tratamiento de datos espaciales en hidrología.* Bubok, 2010.

determinado[87]. Por ejemplo, una base de datos cartográfica se puede construir a partir del plano de una ciudad; en él podemos encontrar entidades reales concretas como los edificios, las vías de comunicación, las plazas, los jardines, las estatuas, o los ríos, y otras abstractas como las divisiones administrativas, los barrios, los distritos etc. Todos estos datos, en formato digital, se pueden almacenar en una base de datos geográfica y de este modo introducirlos en un SIG para su manipulación, pero para que el sistema los entienda se deben ajustar a una estructura de datos según un modelo espacial, que como ya se ha visto puede ser de tipo vectorial o ráster. A su vez, el modelo depende de los parámetros espaciales como el sistema de referencia, es decir, la proyección, el sistema de coordenadas y la escala, y de la resolución de los datos.

En el caso especial del modelo vectorial, junto a los aspectos geométricos de las entidades, es importante tener en cuenta las posibles relaciones espaciales entre éstas, que son almacenadas y administradas en el denominado *modelo topológico,* en el que las entidades se representan por medio de una o varias primitivas topológicas que son: los nodos, los arcos y las caras[88].

Las relaciones topológicas[89] ayudan a conocer la posición relativa de los objetos o entidades sin necesidad de conocer las coordenadas exactas, es decir, nos dan una posición con respecto a otra conocida por la relación espacial que existe entre ambas. Las relaciones espaciales pueden darse entre objetos del mismo tipo o diferentes[90], por ejemplo, en el plano de una ciudad podemos definir la situación de una calle con respecto a los edificios que tiene en sus márgenes.

87. Albert K. W. Yeung y G. Brent Hall, *Spatial database systems: design, implementation and project management.* Springer, 2007 p. (93).

88. 'Desktop Help 10.0 - Fundamentos de topología', 2011 <http://help.arcgis.com/es/arcgisdesktop/10.0/help/#/na/006200000002000000/> [accedido 19 Julio 2011].

89. National Center for Geographic Information & Analysis (U.S.) et al., 'A Framework for the Definition of Topological Relationships and an Algebraic Approach to Spatial Reasoning within this Framework', *National Center for Geographic Information and Analysis*, 1991 <http://www.ugr.es/~cuadgeo/docs/revistas/039.pdf> [accedido 3 Julio 2011].

90. Javier Gutiérrez Puebla y Michael Gould, *SIG: Sistemas de Información Geográfica.* Madrid: Síntesis, 1994 pp. (48-49, 100).

La manera de almacenar las relaciones espaciales entre las primitivas topológicas define diferentes niveles de estructuración topológica[91]:

- *Grafo no planar*, utiliza arcos y nodos, pero en la intersección de dos arcos no se crea un nodo.
- *Grafo planar*, utiliza arcos y nodos, en la intersección de dos arcos se crea un nodo.
- *Topología completa*, se utilizan todas las primitivas topológicas, quedando el espacio cubierto por polígonos sin huecos (*gaps*) ni solapes.
- *Topología abstracta*, no utiliza primitivas geométricas, sólo topológicas, un ejemplo sería el plano del Metro.

En una topología completa se pueden establecer las siguientes relaciones topológicas entre primitivas:

- Los nodos se relacionan con los arcos que confluyen en ellos y con la cara en la que se encuentren si son nodos aislados.
- Los arcos se relacionan con sus nodos inicial y final y con las caras que quedan a su derecha y a su izquierda.
- Las caras se relacionan con los arcos que las delimitan, los nodos aislados que contienen y los arcos interiores a ellas.
- Relaciones entre arcos y nodos.
- Relaciones entre arcos y caras.
- Relaciones entre caras y nodos.

Para entender estas relaciones fijémonos en la Figura 9, en ella aparecen varias entidades espaciales cuya forma geométrica se ha representado por medio de primitivas de dibujo, que a su vez llevan asociadas primitivas topológicas para describir sus relaciones y características espaciales.

---

91. 'The OpenGIS™ Abstract Specification. Topic 1: Feature Geometry (ISO 19107 Spatial Schema). Version 5. OpenGIS™Project Document Number 01-101', ed por OGC John Herring, 2001 <http://www.opengeospatial.org/standards/as> [accedido 19 Junio 2011].

- El arco **5** tiene a **b** como nodo inicial y a **c** como nodo final.
- El arco **4** tiene a su derecha el polígono A y a su izquierda el polígono C.
- El polígono D está incluido en C.
- Los polígonos A y B comparten el arco **5** y los nodos **b** y **c**.
- Los polígonos A, B y C son adyacentes.
- Los arcos **1**, **2**, **3**, **4**, **5** y **6** son contiguos y **4**, **5** y **6** comparten el nodo **b**.

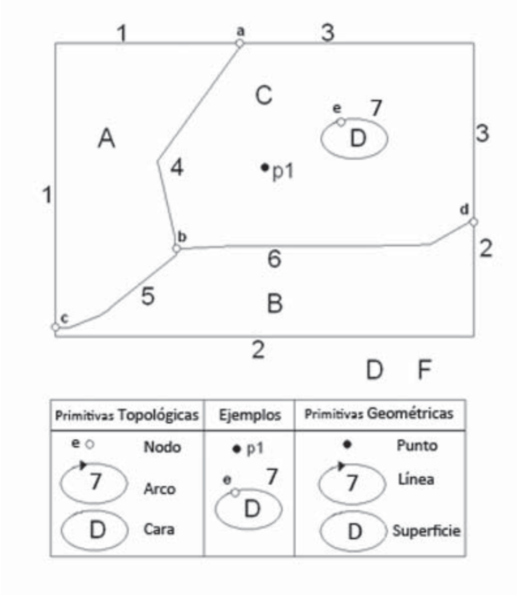


Figura 9. Primitivas y relaciones topológicas.

Un factor importante en el diseño de una base de datos geográfica se refiere a cómo estructurar la información en el sistema para poder acceder a ella de manera eficaz. Por lo general, en un SIG la información se presenta dividida en bloques y/o capas; los *bloques* representan una división horizontal de los datos según criterios espaciales, se realizan para trabajar de forma más detallada y rápida, ya que el volumen de información es menor, por ejemplo, hojas, unidades administrativas,

mallas, etc., mientras que las *capas* representan una división vertical de los datos en función de la temática o geometría de las clases de entidad[92]. Los datos contenidos en las capas o bloques se pueden unir o relacionar para realizar análisis, consultas o ediciones, pero es necesario que exista una armonización espacial y temática, vertical y horizontal, de este modo se consigue una base de datos geográfica continua y homogénea[93].

El siguiente ejemplo es un caso de uso aplicado a las Ciencias Humanas y Sociales. Se trata del diseño de la BDG[94] generada a partir del plano de Madrid realizado por el cartógrafo español Facundo Cañada López en 1902. Después de analizar la información disponible sobre el plano y comprobar los parámetros de exactitud y precisión, se decidió realizar una BDG que vinculara geográficamente la información demográfica de la época con el espacio territorial de Madrid histórico y, de este modo, permitir la realización de diferentes consultas y análisis sociológicos.

Los primeros trabajos se ocuparon de transformar la información analógica en digital escaneando el plano e introduciéndolo en el sistema, después se georreferenció, dotándole de un sistema de referencia y coordenadas que permitieran realizar medidas, cálculo de áreas, etc. Las tareas siguientes se centraron en la elaboración del modelo conceptual, lógico y espacial de la base de datos cartográfica, siguiendo la notación UML en un modelo entidad-relación.

Las entidades espaciales se representaron de acuerdo a un modelo vectorial y se digitalizaron manual y semiautomáticamente, estructurándolas en capas según su geometría y temática; así por ejemplo, con geometría lineal se digitalizaron las capas de calles, las carreteras, los transportes y los cauces de agua; con geometría de polígonos las manzanas y los edificios; y con la de puntos las esculturas ornamentales de la ciudad, o las zonas de interés como puentes, norias, etc. La integridad espacial de la base de datos se logró por medio del modelo topológico

92. Tor Bernhardsen, *op. cit.* p.225.

93. Jordi Guimet Pereña, *Introducción conceptual a los Sistemas de Información Geográfica (S.I.G.).* Madrid: Estudio Gráfico Madrid, D.L., 1992 p. (76).

94. Isabel Del Bosque González et al., *op. cit.*

ajustándose a unas tolerancias adecuadas de acuerdo con la escala, lo que permitirá realizar en un futuro análisis espaciales para estudiar diferentes fenómenos demográficos y sociológicos ocurridos en la época en la ciudad de Madrid.

En conclusión, la información espacial referida a un proyecto puede ser muy heterogénea por el modelo de representación, los métodos y fuentes de adquisición, los sistemas de referencia, la exactitud, la calidad y la posibilidad de actualización, entre otros, pero para que ésta represente la realidad geográfica y pueda ser útil a los fines de un proyecto SIG, tiene que estar integrada y estructurada en una base de datos geográfica, que va a ser, por tanto, el corazón del sistema.

Las Fuentes de Datos

En la actualidad, la información que es capaz de almacenar un SIG admite una gran variedad de formatos y diversidad de fuentes, sin embargo todos estos datos tienen una característica común, su naturaleza digital. Los datos analógicos, como mapas, fotografías aéreas, etc., también pueden ser incluidos en un SIG pero previamente han de ser convertidos a formato digital; éstos por lo general, se utilizan como base para obtener datos nuevos para el SIG.

Los datos se pueden clasificar en *primarios o secundarios* dependiendo de su origen. Los *primarios* se obtienen directamente de la realidad, en un formato adecuado para ser almacenados, procesados y analizados en un SIG, como es el caso de las coordenadas obtenidas a partir de un GPS, las de levantamientos topográficos, prospecciones, etc. Los *secundarios* se obtienen a partir de datos previos, cuyo formato puede ser o no el adecuado para ser introducido directamente en un SIG, como por ejemplo la cartografía impresa[95].

Existen muchas fuentes de información geográfica, pero los principales *métodos de captura de datos* utilizados en un SIG son la digitalización, la fotogrametría, los levantamientos topográficos, el GPS, la teledetección, la información geográfica voluntaria, etc.

95. James Conolly y Mark Lake, *op. cit.*, pp. (95-98).

1. Digitalización

En el proceso de digitalización se genera un fichero digital a partir de un documento analógico. La digitalización puede ser de varios tipos: manual, automática o semiautomática. La elección de una u otra dependerá de la fuente de datos original; por ejemplo, si disponemos de una fotografía aérea antigua en soporte analógico, en primer lugar se realizará una digitalización automática para introducir la información en el SIG en formato ráster, el proceso puede concluir aquí o continuar con otro tipo de digitalización manual o semiautomática que dependerá, fundamentalmente, de las variables con las que se quiera trabajar, los costes económicos (ya que la digitalización manual es más cara), la precisión y la exactitud.

1.1. Digitalización automática

En la digitalización automática, un sistema se encarga de crear los ficheros digitales, de manera automática, a partir de los analógicos. Este tipo de digitalización es la habitual para crear ficheros de tipo ráster, aunque también pueden crearse ficheros vectoriales de forma automática o semiautomática, para lo cual el documento original ha de presentar unas características que permitan el proceso. La digitalización automática se puede realizar mediante *escáner* o por *vectorización automática o semiautomática.*

1.1.1. El escáner

Mediante un mecanismo electro-óptico se puede generar un documento digital con formato ráster a partir de un documento impreso. Un escáner se compone de un cabezal con un sensor que, durante el proceso de escaneo, recorre el documento emitiendo un haz de luz sobre el documento gráfico que la refleja con diferente intensidad, esta reflectancia se recoge generando una matriz de valores; el resultado es un fichero digital de tipo ráster.

Todos los escáneres presentan unas resoluciones radiométricas y espaciales. La *resolución radiométrica* nos informa de la capacidad del

sensor para distinguir entre dos colores distintos. La *resolución espacial* nos indica el número de píxeles por unidad de distancia sobre el papel (dpi)[96]. Es muy importante tener en cuenta la relación entre el tamaño de un píxel (medido en la realidad) y el tamaño de éste medido en la imagen, ya que de esta relación dependerá el volumen de información almacenada, los costes del proceso y la velocidad de acceso a los datos[97].

1.1.2. Vectorización Automática o Semiautomática

Es una técnica que genera ficheros vectoriales partiendo de imágenes digitales en formato ráster, de manera *automática o semiautomática*; en el primer caso se dispone de un software con unos algoritmos capaces de reconocer entidades sobre la imagen, y en el segundo se utiliza un barredor óptico vectorial con un seguidor de línea, en las intersecciones se necesita recurrir a la componente humana para tomar una decisión.

1.2. Digitalización Manual

Consiste en generar un fichero digital a partir de un documento analógico con la intervención humana. El documento obtenido es una capa de tipo vectorial. Con esta técnica sólo se digitalizan entidades reconocibles sobre el documento base. Este proceso es costoso y depende en gran medida del factor humano, pero es el más preciso en cuanto a la representación e interpretación de las entidades, y además permite la corrección de errores durante el proceso.

En la actualidad la digitalización manual se realiza directamente sobre la pantalla del ordenador. Otra técnica, prácticamente en desuso, se basa en el empleo de la tableta digitalizadora.

La *digitalización en pantalla* requiere trabajar sobre un documento base digital, generalmente en formato ráster, procedente de un vuelo aéreo, un satélite, un mapa, otro fichero vectorial, etc. Como en el caso

96. Dpi: dots per inch, puntos por pulgada.
97. Tor Bernhardsen, *op. cit.* pp. (140-141).

anterior, se digitalizan las entidades reconocibles sobre la capa base generando un fichero de tipo vectorial. Este sistema permite la superposición de varias capas cuando se trabaja con un software SIG[98].

1.2.1. Digitalización directa

Para la edición de una capa o entidad geográfica no se necesita previamente una capa base, en ocasiones se cuenta con listados de coordenadas susceptibles de ser introducidos en un SIG. Al proceso de asignar una localización geográfica a puntos de interés, fenómenos, entidades, sucesos, etc., se le conoce como *geocodificación*. La procedencia de estos datos es variable: levantamientos topográficos, coordenadas procedentes de algún fenómeno o suceso, localización espacial, como por ejemplo de los códigos postales, incluso páginas web, GPS, etc. Esta información se introduce directamente en un SIG a través de una tarjeta electrónica, del teclado o importando los datos de hojas de cálculo, de otras bases de datos, de documentos de texto, etc.[99]

2. Fotogrametría

Según la Sociedad Americana de Fotogrametría y Teledetección (ASPR), se define la Fotogrametría[100] como *"el arte, la ciencia y la tecnología capaz de obtener información fiable acerca de objetos físicos y su entorno a través de procesos de grabación, medición e interpretación de imágenes fotográficas y patrones de energía electromagnética radiante y otros fenómenos".*

Es, por tanto, una técnica capaz de suministrar información ráster y vectorial de la superficie terrestre y su posterior integración en un SIG[101]. Por un lado, posibilita la obtención de información métrica bidimensional y tridimensional a partir de las imágenes fotográficas,

98. Paul A Longley et al., *op. cit.,* p. (207).

99. Victor Olaya, *op. cit.,* pp. (125-137).

100. American Society of Photogrammetry. et al., *Manual of photogrammetry.* Falls Church, Va.: American Society of Photogrammetry, 1980.

101. Paul R. Wolf y Bon A. Dewitt, *Elements of photogrammetry: with applications in GIS.* Boston: McGraw-Hill, 2000.

mediante la reconstrucción espacial de los objetos y las superficies registradas, generando imágenes corregidas de distorsión y en proyección (sobre todo en fotogrametría digital)[102], aptas para usos cartográficos, medioambientales, levantamientos arqueológicos, arquitectónicos, etc. Por otro lado, permite la digitalización de las entidades a partir del reconocimiento e identificación de los objetos mediante técnicas de fotointerpretación. Por ello, los tipos de datos geográficos que se pueden obtener por técnicas fotogramétricas y formar parte de las BDG son muy numerosos, como por ejemplo:

- *Datos ráster*: fotografías aéreas, mosaicos, ortofotos, ortofotomapas, ortomosaicos…
- *Datos vectoriales*: mapas topográficos, mapas temáticos, perfiles del terreno, modelados tridimensionales de objetos…
- *Datos numéricos*: coordenadas de puntos de apoyo, del terreno según una cuadrícula, o aleatorios para generar modelos digitales del terreno.

3. Teledetección

Es una técnica que genera imágenes o datos digitales, con diferentes resoluciones, de la superficie de la Tierra por medio de sensores remotos. De acuerdo a la definición de la ASPRS, recogida por Colwell[103] en 1983, la Teledetección puede definirse como *"la medida o adquisición de algunas propiedades de un objeto o fenómeno, y registro de las mismas sin que exista contacto directo con él"*.

La información obtenida por teledetección se corresponde con los valores de radiación que emiten los objetos de la superficie terrestre y que es registrada por *sensores* instalados en plataformas aerotransportadas y satélites artificiales. Los sensores captan la *radiancia* reflejada por los objetos que se encuentran dentro de su campo de visión produciendo una imagen que puede tener múltiples bandas; cada una de ellas, por lo general, representa un segmento del espectro electromagnético

102. José Luis Lerma García, *Fotogrametría moderna: analítica y digital.* Valencia (Spain): Editorial de la UPV, 2002.
103. John R. Jensen, *Remote Sensing of the Environment, an Earth resource perspective.* 2o edn. New Jersey: Prentice Hall, 2000.

registrado por el sensor. Las bandas pueden recoger cualquier porción del espectro, incluidas las series que no son visibles al ojo humano, como es el caso de las secciones de infrarrojos o las microondas.

Dichas imágenes son ficheros ráster de inmediata integración en un SIG, en cuyos píxeles se recoge el valor de la intensidad de la radiación; estos valores son denominados *niveles digitales*[104]. Cuando hay varias bandas, cada píxel con una posición geométrica, posee más de un valor asociado a él. La radiación reflejada por las diferentes superficies de la Tierra, en función de la longitud de onda, origina una curva denominada *firma espectral* de la superficie.

En general, los productos obtenidos por un sistema de teledetección van a depender del tipo de sensor y de su *resolución.*

Se entiende por resolución de un sensor la capacidad de un sistema óptico para distinguir entre señales que están próximas espacialmente o que son espectralmente similares[105]. De acuerdo con la definición anterior, se pueden distinguir cuatro tipos principales de resoluciones asociadas a los distintos sistemas de teledetección: *espacial, espectral, temporal y radiométrica*:

- La *resolución espacial* se refiere a la dimensión real que un píxel de la imagen tiene en su homólogo en el terreno.
- La *resolución espectral* viene dada por la región del espectro electromagnético y el número de bandas en las que un sensor registra esta información.
- La *resolución temporal* se refiere al tiempo que tarda un sensor en registrar datos de una misma zona.
- La *resolución radiométrica* es la capacidad del sensor para detectar variaciones en la radiación espectral.

A su vez, los sensores se suelen clasificar en *pasivos* y *activos*, dependiendo del origen de la radiación; los primeros recogen la energía radiante reflejada por los objetos de la superficie de la Tierra procedente del Sol. Estos a su vez, dependiendo de cómo registren la información en las imágenes, pueden ser de *barrido,* que registran valores de radia-

104. Paul J. Gibson y Clare H. Power, *Introductory Remote Sensing. Digital Image Processing and Applications.* Taylor & Francis, 2000 pp. (2-7).
105. John R. Jensen, *op. cit.,* p. (12).

ción píxel a píxel, o de *empuje,* que la recogen línea a línea. Los activos emiten radiación y la recogen tras ser reflejada por los objetos; de este tipo son los datos RADAR y LIDAR.

La tecnología RADAR[106] (*Radio Detection and Ranging*) emite pulsos de ondas de radio y recoge la reflexión sobre la superficie observada en un sensor. Una de las aplicaciones más importante ha sido la generación de un modelo digital de elevación (MDE, DEM en inglés) a nivel global mediante procesos de interferometría RADAR desde el espacio: el SRTM (*Shuttle Radar Topography Mission*), cuya precisión en vertical está entre 6 y 16 metros y posee una resolución espacial entre 30 y 90 metros[107].

En el LIDAR (*Light Intensity Detection and Ranging*) se utilizan pulsos de luz polarizada —entre el ultravioleta y el infrarrojo próximo— mediante un emisor láser, proporcionando varios retornos de la energía que son recogidos por sensores aerotransportados; esta particularidad, junto a la consolidación actual de esta tecnología a nivel mundial, hace que los datos LIDAR sean muy utilizados para realizar cartografía de elevaciones, por ejemplo; modelos de contorno y MDE para ortofotos[108].

3.1. Características de los satélites y sensores más utilizados en un SIG

Directamente integrables en los sistemas de tratamiento de imágenes, de información geográfica o de redacción cartográfica, las imágenes de satélite pueden procesarse y asociarse fácilmente con otros datos geográficos para extraer toda la información útil. Por ello, cada vez existe un mayor número de programas y sensores que proporcionan información de la superficie terrestre. Algunos de los más comunes son LANDSAT, SPOT, IKONOS, QUICKBIRD, ASTER, NOAA-AVHRR

106. Natural Resources Canada. Canada Centre for Remote Sensing, 'Introducción a la percepción remota mediante radar'. <http://www.ccrs.nrcan.gc.ca/resource/tutor/gsarcd/pdf/bas_intro_s.pdf> [accedido 11 Julio 2011].

107. Para más información, se puede visitar la página de la NASA: http://www2.jpl.nasa.gov/srtm/

108. Karl Kraus y Norbert Pfeifer, 'Advanced DTM generation from Lidar data', *International Archives of Photogrammetry and Remote Sensing*, XXXIV-3/W4 Annapolis, MD, (2001) pp. (23-30).

o MODIS. Sin pretender realizar una relación exhaustiva de los mismos, en la Tabla 2 se detallan las principales características[109] de algunos de los más utilizados en SIG.

SATÉLITE	SENSOR	RESOLUCIÓN			ANCHO BARRIDO
		ESPECTRAL	ESPACIAL	TEMPORAL	(Km)
LANDSAT5	TM	$B_{1,2,3,4,5y7}$	30 m	16 días	185
		B_{6Ly6H}	120 m		
LANDSAT7	ETM+	$B._{1,2,3,4,5y7}$	30 m		
		$B._{6Ly6H\ (IRT)}$	60 m		
		1 Pan	15 m		
SPOT5	HRG	2 Pan	2,5m y 5m	26 días	60
		3B.multiespectrales	10 m		
		1IRM	20 m		
	VEGETA-TION2	4 Bandas: Azul, Rojo, IRP, IRM	1 km	1 día	2250
NOAA	AVHRR	6 Bandas del verde al IRT	1 km	6 horas	2900
QUICKBIRD	BGIS2000	Pan	0,62 a 0,72m	1 a 3,5 días	16,5
		Azul, Verde, Roja, IRP	2,44 a 2,88m		
TERRA /AQUA	MODIS	$B_{1,2}$	250 m	Diario	2330
		B_{3-7}	500 m		
		B_{8-36}	1000 m		
	ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer)	$IRP_{B\ 1,2,3}$	15 m	16 días	60
		$IRM_{B\ 4-9}$	30 m		
		$IRT_{B\ 10-14}$	90 m		
IKONOS	IKONOS	Multiespectrales	3,2 m	3 a 5 días	11
		Pan	0,82 m		
		Pan	0,82 m		

Tabla 2. Sensores e imágenes de satélite.

109. 'Sensors Used in Satellite Imaging | Satellite Imaging Corp', 201 <http://www.satimagingcorp.com/satellite-sensors.html> [accedido 18 Mayo 2011].

Las *principales aplicaciones* de estos sensores son las que se detallan a continuación:

- *LANDSAT5 y LANDSAT7:* para detección de cambios antropogénicos y naturales (desarrollo agrícola, deforestación, desertificación, etc.).Una muy significativa es la aplicación en el proyecto europeo CORINE Land Cover 1990 y 2000[110].
- *SPOT5:* óptimo para cartografías de escalas medias (1:25.000 a 1:10.000), como en el proyecto SIOSE 2006[111]. Produce imágenes estéreo par utilizadas en MDE, simuladores de vuelo o planificación de telefonía móvil, por ejemplo.
- *NOAA:* para detección de cambios en las coberturas terrestres globales.
- *QUICKBIRD:* para cambio de usos del suelo, exploración de gas, cartografía de gran escala, control de áreas costeras sensibles, modelado en 3D, etc.[112]
- *TERRA/AQUA de MODIS:* para detectar nubes, aerosoles, vapor atmosférico, masas de agua, temperatura atmosférica y superficial, riesgo de incendios forestales, etc.
- *TERRA/AQUA de ASTER:* para mapas en detalle de temperatura de la superficie terrestre, mapas de reflectancia y MDE gracias a la visión estereoscópica, como el ASTER GDEM[113].
- *IKONOS*: se utiliza con frecuencia en aplicaciones medioambientales, en cartografía urbana y rural, recursos naturales, agricultura, ingeniería, minería, construcción, detección de cambios, monitoreo de costas, análisis de superficie en 3D y MDE por su posibilidad estereoscópica.

---

110. Antonio Arozarena Villar et al., *Mapa de Ocupación del Suelo en España. Corine Land Cover-Proyecto I&CLC2000.* Madrid: Centro Nacional de Información Geográfica, 2006 <http://hdl.handle.net/10261/26004> [accedido 27 Julio 2011].

111. Véase http://www.ign.es/ign/layoutIn/coberturaUsoSuelo.do.

112. Marta Yebra, Emilio Chuvieco y David Riaño, 'Estimation of live fuel moisture content from MODIS images for fire risk assessment', *Agricultural and Forest Meteorology*, 148 (2008) pp. (523-536).

113. El ASTER GDEM es un modelo digital de elevaciones global, con una resolución espacial de 1 arco de segundo, unos 30 m en planimetría y 20m en altitud. Se ha obtenido mediante estereo-correlación automática a partir de las imágenes ASTER. Posee una cobertura del 99% de la superficie terrestre: Más detalles en la página oficial, http://www.gdem.aster.ersdac.or.jp/

4. Plan Nacional de Observación del Territorio (PNOT)

El Instituto Geográfico Nacional (IGN)[114] es el organismo oficial con competencias en materia de información geográfica en España. Es por ello, que es el mayor proveedor de datos geográficos de nuestro territorio, poniendo a disposición de toda la comunidad científica y técnica y del ciudadano en general datos básicos y de referencia, a nivel nacional, de alto grado de detalle y actualizados periódicamente.

Gracias a proyectos colaborativos, como el Plan Nacional de Observación del Territorio (PNOT)[115], liderado por el IGN, y en el que participan otros organismos de la Administración General del Estado (AGE), a través de diversos departamentos ministeriales, y todas las Comunidades Autónomas (CCAA), se han aunado los esfuerzos técnicos, logísticos y económicos necesarios para la adquisición, tratamiento y extracción de información geográfica a partir de imágenes aeroespaciales. Este proyecto pretende optimizar los recursos de las administraciones públicas, evitando la duplicidad de esfuerzos y asegurando la armonización de las bases de datos geográficas y la normalización de los procesos. Bajo el paraguas de PNOT se han adquirido, procesado y puesto a disposición, una gama amplia de productos geográficos en España desde el año 2004 (coberturas completas de imágenes aéreas, de satélite, modelos digitales de elevaciones, etc.), con el objetivo general de que sean utilizadas en aplicaciones multidisciplinares, con resoluciones y periodos de actualización diversos. Este macro-proyecto se materializa en planes específicos[116] concretos:

- *Plan Nacional de Ortofotografía Aérea (PNOA)*, que proporciona, entre otros, productos fotogramétricos, coberturas periódicas (de dos a cuatro años) de todo el territorio nacional mediante ortofotografía aérea de alta y muy alta resolución, PNOA25/50 (25/50 cm) y PNOA10 (10 cm).

114. La URL de la página web del IGN es: http://www.ign.es/ign/main/index.do.

115. Antonio Arozarena et al., ‘El Plan Nacional de Observación del Territorio en España: Situación actual y próximos pasos.’, *Mapping Interactivo*, 2006 <http://www.mappinginteractivo.com/plantilla-ante.asp?id_articulo=1206> [accedido 1 Septiembre 2011].

116. Para más información: http://www.ign.es/Memoria/ES/AcCaptacion.html.

- *Plan Nacional de Teledetección (PNT)*, que proporciona coberturas periódicas (anuales, mensuales y semanales) de todo el territorio nacional mediante imágenes de satélite de media y baja resolución (2,5 a 100 m).
- *Sistema de Información de Ocupación del Suelo en España (SIOSE),* que proporciona información de coberturas y usos del suelo a escalas medias (1:25000), con un modelo de datos multi-parámetro, multi-criterio e integrador de la información de las bases de datos geográficas de las CCAA y la AGE, y con un periodo de actualización de 5 años.

5. Levantamientos topográficos

Proporcionan una relación de coordenadas geográficas, las cuales pueden ser introducidas directamente en un SIG, que permiten la representación de un fenómeno del territorio o de un área de la superficie terrestre.

Un levantamiento topográfico comprende las operaciones y medidas necesarias para la obtención en campo de los datos que permitan la representación de un territorio en un mapa o un plano, de acuerdo con una escala determinada[117]. Los levantamientos topográficos suelen realizarse sobre un área de la Tierra relativamente pequeña, donde la curvatura de la misma va a ser despreciable.

En los trabajos de campo, se miden ángulos y distancias desde puntos de coordenadas conocidas de una red topográfica, generalmente por el método de radiación, a los puntos cuyas coordenadas queremos determinar. En la actualidad, para la realización de estas medidas se utilizan aparatos de topografía muy precisos, como las *Estaciones Totales* y los *GPS topográficos* de corrección diferencial. A partir de las medidas tomadas en campo, y mediante cálculos trigonométricos, se determinan las coordenadas que permitirán el trazado de los objetos en un plano, un mapa, o un modelo digital de elevaciones.

117. Francisco Domínguez, *Topografía general y aplicada.* Barcelona: Mundi-Prensa, 1986.

5.1. El láser escáner 3d

El mundo de la topografía ha evolucionado considerablemente en los últimos años, con la aparición de nuevos aparatos de medida que permiten el registro de los objetos de una manera cada vez más fiable, o con la captación de un gran volumen de datos y complejas técnicas de procesado que posibilitan su posterior representación de una manera precisa, y su integración posterior en un SIG junto con otro tipo de datos territoriales.

La reciente aparición del Láser Escáner 3D (también denominado *Láser Escáner Terrestre*, para diferenciarlo de otras técnicas de captura de datos mediante láser como el LIDAR, mencionadas en este mismo capítulo), están permitiendo la reconstrucción de espacios tridimensionales, como los relativos a los espacios naturales, o los registros geométricos de entidades de patrimonio cultural[118], de una manera eficiente, fiable, precisa y complementaria con los métodos topográficos y fotogramétricos tradicionales.

El láser escáner terrestre es un instrumento de medida que realiza un barrido de las superficies mediante un haz de luz láser, creando una densa nube de puntos (x, y, z) a partir de la cual se pueden reconstruir, con gran calidad y alta resolución, la forma de los objetos y los colores de su superficie. La utilización del láser escáner y la elaboración de modelos digitales 3D es muy utilizado, por ejemplo, en el diseño y prototipos industriales, análisis de estructuras, documentación métrica en arqueología y arquitectura[119], etc.

118. Mercedes Farjas Abadía y Francisco J. García Lázaro, *Modelización tridimensional y sistemas láser aplicados al Patrimonio Histórico*, Biblioteca básica. La Ergástula, 2008.

119. Patricia Mañana-Borrazás y Anxo Rodríguez Paz, 'Una experiencia en la aplicación del Láser Escáner 3D a los procesos de documentación y análisis del Patrimonio Construido: su aplicación a Santa Eulalia de Bóveda (Lugo) y San Fiz de Solovio (Santiago de Compostela)', *Application of Terrestrial Laser Scanner to recording and analysis of Building Heritage: cases study of Santa Eulalia de Bóveda (Lugo) and San Fiz de Solovio (Santiago de Compostela)*, 2008 <http://hdl.handle.net/10261/15849> [accedido 2 Septiembre 2011].

6. Sistema de Posicionamiento Global (GPS)

Como en el caso de los levantamientos topográficos, proporciona un listado de coordenadas que pueden ser introducidas directamente en un SIG.

GPS es un sistema de radionavegación estadounidense basado en técnicas espaciales, que posibilita la determinación de la localización y la hora local de los puntos de la superficie terrestre, con criterios de precisión. El sistema se compone de tres elementos básicos: una constelación de 24 satélites de órbita conocida alrededor de la Tierra, las estaciones terrestres de seguimiento y control que introducen las correcciones necesarias para mantenerlos en la órbita apropiada, y los receptores GPS de los propios usuarios.

Desde el espacio, los satélites transmiten señales de radio que reciben e identifican los receptores GPS; éstos, a su vez, calculan por separado las coordenadas de los puntos (X, Y, Z), así como la hora local precisa. La posición de los satélites es conocida a través de las efemérides orbitales y en base a estos parámetros, el aparato GPS puede calcular las coordenadas de cualquier punto a través de un proceso de *trilateración*, basado en el conocimiento de las distancias a partir de las señales emitidas y el tiempo transcurrido.

Estos receptores, permiten determinar localizaciones con precisiones subcentrimétricas[120], así como el cálculo de rutas. Las aplicaciones son muy numerosas, tanto en el ámbito militar como en el civil, sirviendo de apoyo a la navegación aérea, terrestre y marítima, los servicios de emergencia y socorro, la telefonía móvil, las redes de distribución eléctrica, la agricultura, el medioambiente, la cartografía y la geodesia, la geología, el ocio, etc.

Este sistema proporciona información geográfica de un modo rápido, eficaz, seguro, preciso y económico, ya que las señales del GPS son gratuitas, aunque al ser el sistema propiedad del gobierno ameri-

120. 'Welcome to GPS.gov', 2011 <http://www.gps.gov/> [accedido 18 Mayo 2011].

cano están sujetas a disponibilidad selectiva, como en el caso de los recientes conflictos armados ocurridos en Oriente Medio, por ejemplo.

7. Importación directa de Datos

Es otra fuente de datos que incorpora dentro de los SIG información alfanumérica diversa existente en otras bases de datos, hojas de cálculo u otros formatos, de este modo se puede completar la información geográfica del proyecto. En algunos casos, podría ser necesario un cambio previo de formato que posibilite su integración.

8. Información Geográfica Voluntaria (VGI)

Puede considerarse como una forma nueva no convencional de generar o actualizar datos espaciales con la participación de individuos que, de modo propio y gratuito, ofrecen información geográfica a través de la Web. Según M.F.Goodchild[121] se conoce como VGI al uso de internet para crear, gestionar y difundir información geográfica aportada voluntariamente por usuarios de la propia red. La información procedente de VGI tiene una calidad inferior a la obtenida de manera clásica, sin embargo su precisión y validez vienen determinadas por la metodología e instrumental utilizados en la creación y obtención del dato.

121. Michael F. Goodchild, 'Citizens as sensors: the world of volunteered geography', *GeoJournal*, 69 (2007) pp. (211-221).

Capítulo IV
Edición y consulta de datos

En los capítulos anteriores hemos ido siguiendo las fases de un proyecto SIG, en el punto en que nos encontramos y, suponiendo que hemos seguido dichas fases, deberíamos tener los datos procedentes de diversas fuentes almacenados en sus correspondientes bases de datos geográficas, pero antes de seguir avanzando debemos comprobar que no existen errores. Si los hubiera tendríamos que corregirlos con las herramientas de edición temática y espacial que nos ofrecen los SIG, ya que de la calidad de nuestra información dependerá la validez de las actualizaciones, o las consultas, los análisis y la representación de la información posterior.

Edición de datos

Cuando se cargan los datos geográficos en un SIG deben ser tratados para adecuarlos al modelo y estructura del proyecto, así como para corregir los errores que contengan, ya sean de índole espacial o temática, para lo cual son sometidos a un proceso de edición, en el que además se pueden crear nuevos datos. Los errores que a menudo presentan los datos son debidos a[122,123]:

- El factor humano, que puede originar en el proceso de digitalización errores geométricos, topológicos, de interpretación de la realidad y semánticos.

122. José Miguel Santos Preciado y Francisco Javier García Lázaro, *Análisis estadístico de la información geográfica*, 1o edn. Madrid: Librería UNED, 2008 p. (50).
123. Tor Bernhardsen, *op. cit.,* pp. (205-206).

- El propio proceso de carga de los datos en el sistema.
- Los originados en el proceso de captura.
- La integración de datos heterogéneos en el sistema con una deficiente armonización horizontal o vertical.

El proceso de edición de los datos geográficos se contempla desde sus dos componentes principales, la *temática* y la *espacial*, y el tipo de información sobre la que se realiza puede ser vectorial o ráster, siendo más frecuente la primera.

1. Edición Temática

La componen todas aquellas tareas relacionadas con la creación y corrección de los aspectos semánticos o atributos de las entidades, para su correcta identificación y situación espacial.

En las tareas básicas de edición temática los SIG utilizan herramientas estándar como las de cualquier procesador de textos para copiar, cortar, pegar, mover, o renombrar; pero cuando las tareas requieren funciones más complejas se valen de lenguajes como SQL, para manipular los atributos de las bases de datos relacionales, así como comandos específicos de las aplicaciones SIG[124].

Las principales *tareas de edición temática* se ocupan de:

- Definir y crear correctamente los atributos de las entidades.
- Establecer los dominios o conjuntos de todos los valores que pueden tomar los atributos.
- Comprobar las etiquetas de las entidades para que no contengan errores semánticos, ni duplicidades o falta de información.
- Añadir atributos a las entidades nuevas, creadas durante un proceso de edición.
- Crear nuevos atributos a partir de los existentes o de los de otras capas.

124. Tor Bernhardsen, *op. cit.,* pp. (233).

2. Edición Geométrica

Se refiere a la edición de la componente espacial de los datos y con ella se pretende generar una base de datos precisa, íntegra y con la calidad requerida por el sistema, de tal forma que no contenga errores espaciales, ni falta de información y que permita su actualización a partir de los datos existentes u otros nuevos.

La edición geométrica puede realizarse atendiendo sólo a los aspectos geométricos o teniendo también en cuenta las relaciones espaciales entre entidades, en cuyo caso hablamos de una *edición topológica*, que permite crear o modificar entidades conforme a unas reglas espaciales[125] y tolerancias previamente fijadas, por ejemplo, en una capa lineal se puede aplicar la siguiente regla: que en las intersecciones las líneas se corten creando nodos; otro ejemplo, en el caso de una capa de polígonos sería que no existieran agujeros o zonas sin información (*gap*) entre ellos; una regla más muy común y aplicable a ambos tipos geométricos es que no existan solapes entre las entidades de una misma capa. Las dos últimas reglas citadas se dan en una topología completa. Existen más reglas y establecerlas ante una edición topológica depende del proyecto.

El otro concepto que se ha nombrado y que tiene una gran importancia a la hora de mantener la coherencia espacial es el de *tolerancia de edición*, que se puede definir como la distancia entre dos puntos del espacio para la cual, o por debajo de la cual, se considera que ambos ocupan el mismo lugar. La tolerancia establecida para una capa con un tipo geométrico ha de mantenerse durante todo el proceso de edición ya que si no se podrían crear zonas sin información, así por ejemplo, si en el transcurso de dos procesos de edición consecutivos en una capa de polígonos se varían las tolerancias puede ocurrir que entre las caras de dos polígonos contiguos se generen microespacios sin información o agujeros o por el contrario solapes.

125. ESRI, 'Desktop Help 10.0 - Reglas topológicas de las geodatabases y soluciones a los errores de topología', 2011 <http://help.arcgis.com/es/arcgisdesktop/10.0/help/index.html#//001t000000sp000000> [accedido 5 Septiembre 2011].

Una de las tareas de edición más importantes se refiere a la creación de nuevas entidades partiendo de la información procedente de una capa base o los valores temáticos de entidades contenidos en una base de datos, utilizando funciones básicas de edición geométrica como copiar, cortar, pegar, borrar, mover, girar, partir, o unir, que también nos permiten modificar las existentes para actualizarlas o corregirlas. Además se pueden realizar otro tipo de tareas más complejas como son los ajustes espaciales[126], a los que deben someterse los bloques o unidades de trabajo que se unen horizontal o verticalmente y que en ocasiones conllevan una adecuación de la escala que se consigue por medio de procesos de generalización, siempre y cuando se pase de una mayor a otra menor entre valores relativamente próximos.

En todos estos casos se producen *errores geométricos o topológicos* (ver Figura 10), siendo los más comunes[127,128]:

- *Arcos colgantes,* son arcos en los que alguno de los nodos de sus extremos no *intersecta* con el nodo de otro arco, originándose un *agujero*, un *sobrepaso* o un *subtrazo*.
- *Falsos polígonos,* también llamados *slivers*, tienen un área muy por debajo de las especificaciones geométricas del proyecto, y se suelen generar al unir capas en sentido vertical u horizontal o en el transcurso rutinario de un proceso de de edición.
- *Bucles*, es una especie de rizadura producida en un arco que genera un mini polígono.
- *Líneas duplicadas*, son partes de una línea o polígono que ocupan el mismo lugar en una misma capa.
- *Vértices desalineados*, son picos de sierra aislados o no, sobre arcos.

La edición correcta de la componente temática y la geométrica con topología es uno de los factores que contribuye a mejorar la calidad de los datos en un SIG, porque asegura la integridad espacial de la

---

126. Donna J. Peuquet y Duane Francis Marble, *Introductory readings in geographic information systems.* Taylor & Francis, 1990 p. (41).
127. Peter A. Burrough y Rachael A. McDonnell, *op. cit.,* pp. (60-63).
128. Donna J. Peuquet y Duane Francis Marble, *op. cit.,* p. (239).

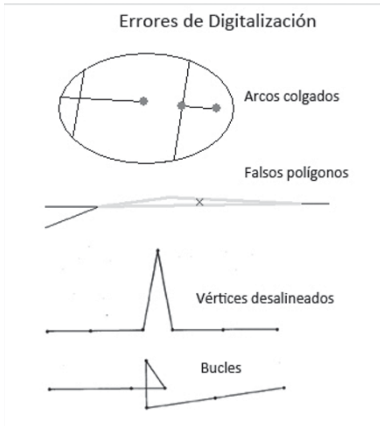


Figura 10. Errores de digitalización

información geográfica, permitiendo de este modo la realización de mediciones precisas, diferentes tipos de cálculos, consultas y numerosos y diversos análisis espaciales; en definitiva, nos permite explotar un SIG con la suficiente seguridad en los resultados.

Generalización

Es un proceso que se aplica a los datos para reducir el volumen de información, mejorando el aspecto general de los mismos y reduciendo el tiempo de procesado.

En esta primera parte del capítulo dedicada a la edición sólo nos vamos a ocupar de las operaciones más frecuentes a las que puede someterse la geometría de las entidades cuando se les aplica un proceso de generalización, los aspectos referentes a la generalización semántica y simbología serán tratados posteriormente en el capítulo dedicado a la representación de los datos.

En la creación y mantenimiento de las bases de datos son necesarios los procesos de generalización, así como para la generación de cartografía a diferentes escalas. En muchas ocasiones este proceso se utiliza para mejorar la estética de las entidades tras un proceso de edi-

ción, o para resaltar determinadas características de las entidades de una determinada zona geográfica, sin que ello conlleve un cambio en la escala.

Las operaciones gráficas a las que se someten las entidades en un proceso de generalización fueron definidas por Robert B. McMaster y K. Stuart Shea [129], y son: simplificación, suavizado, agregación, exageración, desplazamientos, amalgamación, fusión, colapso, refinamiento y mejoras. Desde su definición se han creado diversos algoritmos para aplicar estos cambios, a continuación veremos con algo más de detalle las dos primeras pues son los más habituales en la generalización de la entidades sin cambio de escala.

La *simplificación de líneas* es una de las operaciones más utilizadas en un proceso de generalización, en ella se eliminan algunos vértices de las líneas que no alteran en general la forma original. Para realizarla se han desarrollado diversos algoritmos que se pueden clasificar en[130]:

- Vecindad inmediata, vecindad acotada y vecindad no acotada: en los tres se analiza cada punto y los de su alrededor.
- Globales: en estos se analiza la línea de forma global.
- Los que preservan la topología: en este caso además de atender a las propiedades geométricas locales y globales, se fija en las relaciones espaciales de las líneas para preservarlas.

El algoritmo de simplificación de uso común en cartografía, y que pertenece al grupo de los globales, fue desarrollado por Douglas y Peucker en 1973[131]; este método elimina puntos de una línea basándose

129. Robert Brainerd McMaster y K. Stuart Shea, *Generalization in Cartography-* Washington, DC: Association of American Geographers, 1992 <http://geoinformatics.ntua.gr/courses/admcarto/lecture_notes/generalisation/bibliography/mcmaster_shea_1992.pdf> [accedido 5 Septiembre 2011].

130. Robert Brainerd McMaster, ‘Automated Line Generalization’, *Cartographica: The International Journal for Geographic Information and Geovisualization*, 24(2) (1987) pp. (74-111).

131. David H. Douglas y Thomas K. Peucker, ‘Algorithms for the reduction of the number of points required to represent a digitized line or its caricature’, *Cartographica: The International Journal for Geographic Information and Geovisualization. University of Toronto Press*, 10(2) (1973) pp. (112-122).

en una tolerancia preestablecida: considerados dos puntos de una línea, se traza una línea recta entre ellos y posteriormente se lanzan perpendiculares a esta recta desde diferentes puntos de la línea comprendidos entre los dos primeros; aquél punto cuya distancia a la línea recta sea inferior al valor de tolerancia, se elimina. El proceso se va repitiendo hasta recorrer la longitud de toda línea, al final del proceso se habrá reducido el número total de vértices, sin variar sustancialmente la forma de la misma.

El suavizado, con él se consigue redondear elementos lineales que tras un proceso de digitalización, vectorización automática o por un cambio de escala, presentan un aspecto muy anguloso. Con este proceso también se debe tener cuidado para no modificar la línea, de tal modo que deje de representar la realidad[132]. Como en el caso anterior existen diversos algoritmos de suavizado[133]:

- Media entre puntos.
- Ajuste de funciones matemáticas como *splines* o curvas Bézier.
- Tolerancias.

Estos dos procesos se aplicaron para mejorar el aspecto general de los ejes de calle digitalizados semiautomáticamente a partir del plano de Facundo Cañada de 1902[134]; una vez terminado el proceso de edición se comprobó que la estética general no era muy buena y además se habían creado demasiados vértices. Con una simplificación de líneas se redujo el número de éstos y con un suavizado se consiguió un aspecto más real en las calles. Estos procesos originaron cambios en la base de datos a nivel geométrico y semántico sin un cambio de escala.

Todas estas operaciones geométricas, forman parte de las técnicas empleadas en los procesos de creación de cartografía a partir de un proceso de generalización por *reducción de escala cartográfica* (cuando

132. Victor Olaya, *op. cit.,* pp. (462-465).

133. Robert Brainerd McMaster, 'The integration of simplification and smoothing algorithms in line generalization', *Cartographica: The International Journal for Geographic Information and Geovisualization. University of Toronto Press*, 26(1) (1989) pp. (101-121).

134. Isabel Del Bosque González et al., *op. cit*.

pasamos de una representación detallada de la realidad a otra menor). En este caso las simplificaciones son más drásticas, afectando a entidades que pueden desaparecer en el proceso. En este caso las operaciones más empleadas son[135]:

- La reducción o pérdida de detalle en las líneas y contornos de los polígonos. Por ejemplo, los detalles de la línea de la costa como las bahías pequeñas representadas en un mapa de gran escala pueden perderse al generalizar y suavizar el contorno general de la costa.
- La desaparición de fronteras o arcos entre entidades adyacentes del mismo tipo. Por ejemplo, en un mapa de ocupación del suelo se podrían agrupar polígonos adyacentes cuya unidad jerárquica superior fuera igual. Imaginemos una zona urbana en la que se hubieran diferenciado dos polígonos contiguos, uno con tejido urbano continuo (casco) y otro con tejido urbano discontinuo (ensanche), tras un proceso de generalización ambos podrían unirse eliminando la frontera o cara entre ellos para constituir uno sólo de tejido urbano.
- La unión de unidades u hojas de trabajo por sus bordes. Un ejemplo real lo tenemos en la generalización del MTN25 (Mapa Topográfico Nacional a escala 1:25.000) para obtener el MTN50 (Mapa Topográfico Nacional a escala 1:50.000)[136]. En este caso varias hojas del mapa de mayor detalle formarán una del generalizado.

Como es lógico, en todos estos procesos se aplican las herramientas de edición geométrica y temática ya que, por ejemplo, la eliminación de un arco supone modificar objetos iniciales cuyos atributos también deben ser modificados para adecuarlos a la nueva situación.

135. Donna J. Peuquet y Duane Francis Marble, *op. cit.,* pp. (42-43).

136. Javier Iribas Cardona, 'Obtención del MTN50 por generalización cartográfica del MTN25 digital', *Mapping Interactivo*, 1997.

<http://www.mappinginteractivo.com/plantilla-ante.asp?id_articulo=774> [accedido 14 Julio 2011].

CONSULTAS EN UN SIG

Como vimos en capítulos anteriores, en un SIG la información geográfica se almacena en bases de datos que pueden ser interrogadas por el sistema para seleccionar aquellas entidades que cumplen unas determinadas condiciones; a este proceso se le denomina *consulta.*

La componente temática y espacial de las entidades que componen la información geográfica en un SIG permiten que se puedan realizar consultas utilizando criterios de selección temáticos, espaciales o ambos, en cuyo caso se habla de consultas mixtas. En cada una de ellas se utilizará un lenguaje con los criterios apropiados.

La *sintaxis* en un lenguaje de consulta suele ser[137]:

Selecciona SUJETO CONDICIONES

El *Sujeto* se refiere a las clases de objetos que pueden ser seleccionadas.

Las *Condiciones* son los criterios temáticos y/o espaciales.

A continuación se van a formular una serie de consultas a modo de ejemplo, sobre las entidades de la base de datos cartográfica del Municipio de Madrid descrita en el capítulo III.

1. Selecciona las *calles_cuyo nombre sea* Alcalá (= Alcalá)
   El criterio es temático.

2. Selecciona los *barrios incluidos en* el distrito Palacio
   El criterio es espacial.

3. Selecciona las *calles cuyo nombre empiece por la letra A* y que estén *incluidas en el barrio Argüelles*
   El criterio es temático y espacial.

137. Max J. EGENHOFER, 'Spatial SQL: A Query and Presentation Language', *National Center for Geographic Information and Analysis*, 1994 pp. (86-94).

1. Consultas Temáticas

En este tipo de consultas se realiza una selección de las entidades de una base de datos, basándose en una o varias características de sus atributos.

Los SIG se valen de un lenguaje como el SQL[138] para realizar consultas temáticas en bases de datos (BD) relacionales. La mayoría presentan una interfaz de consulta preparada, siguiendo la sintaxis de este lenguaje, para facilitar las consultas a los usuarios no expertos en él.

Los criterios de selección que se utilizan son operadores aritméticos y/o lógicos, y/o álgebra Booleana[139].

Los *operadores aritméticos* más habituales son:

- +, -, x, /, etc.

Los *operadores lógicos*:

- Mayor que (>), menor que (<), igual (=), mayor o igual que (≥), menor o igual que (≤), diferente a (<>).

Por ejemplo:

Selecciona las manzanas de Madrid cuya superficie sea ***menor o igual*** a 20 metros cuadrados.

Los *operadores Booleanos:*

- AND, OR, NOT, NOR, LIKE, etc.

Por ejemplo:

Selecciona las parcelas de olivos ***y*** las de viñedos del CORINE90.

138. Gregorio Quintana et al., *Aprende SQL.* Castelló de la Plana: Universitat Jaume I, 2008.

139. Javier Gutiérrez Puebla y Michael Gould, *op. cit.,* (172-173).

Las *consultas temáticas* no sólo se realizan entre atributos de una misma clase de objetos o tabla en una BD relacional, también se pueden realizar sobre los atributos de dos o más clases. Esto es posible si previamente se unen sus tablas por medio de algún campo definido de la misma manera y con atributos comunes, de este modo cada registro de una clase (tabla) se asocia con otro u otros de la otra clase. La relación entre los registros de las tablas será 1:1 o 1: M y, como ya vimos en capítulos anteriores, a esta relación se la denomina cardinalidad.

El propósito final de las consultas temáticas es la elaboración de mapas temáticos, en cuya leyenda se muestran los valores de la variable o variables sobre las que se ha hecho la consulta o selección.

En el caso de la información de tipo ráster, estas consultas nos devuelven un conjunto de píxeles que ocupan una determinada posición y cuyos valores temáticos se ajustan a una consulta según una condición aritmética y/o lógica. Esta operación se realiza por la combinación de las operaciones de análisis local: reclasificación y superposición de mapas. Las consultas con condiciones aritméticas nos permiten elaborar mapas cuyos píxeles cumplen estas condiciones, después pueden superponerse todos los mapas utilizando la lógica booleana para obtener un único mapa que nos muestre todas las condiciones a la vez.

Otra forma de realizar consultas en estos sistemas consiste en asociarles una base de datos temática, de tal forma que la consulta nos responderá con los valores espaciales que cumplan las condiciones de los operadores tal y como sucede en los modelos vectoriales.

2. Consultas Espaciales

Este tipo de consultas nos devuelven entidades cuya condición es la de ocupar un determinado lugar establecido por el usuario, o bien presentar una determinada relación espacial con respecto a otra u otras entidades, por lo que se pueden clasificar como:

- Consultas por ámbito espacial.
- Consultas de relaciones espaciales.

La expresión que define estas consultas se diferencia de las temáticas en el operador, que es de tipo espacial, y hace referencia a una determinada situación espacial entre objetos.

2.1. Consultas por Ámbito Espacial

Con este tipo de consultas se seleccionan aquellas entidades situadas en una zona geográfica determinada por el usuario. El elemento espacial de selección puede ser un punto, una línea o un recinto poligonal regular o no, cuyo tamaño sea tan grande como queramos —incluso abarcando todo el proyecto—, en todos los casos la selección de las entidades se realiza por su posición con respecto al elemento de selección. Existen diversos criterios para establecer las condiciones de la consulta que se refieren a las diferentes posiciones de los objetos con respecto a los que se utilicen para realizar la selección, así los objetos pueden ocupar las siguientes posiciones:

- Los objetos espaciales pueden estar completamente dentro de…
- Los objetos espaciales que intersecten con el recinto...
- Los objetos espaciales pueden tocar el recinto o estar completamente dentro de...
- Los objetos que quedan completamente fuera de…
- Los objetos que quedan completamente fuera de o tocan el recinto…

Al realizar la selección gráfica de los objetos quedan seleccionadas las propiedades geométricas y espaciales de éstos como son el área, la longitud, la topología, etc., así como sus registros en la base datos, es decir, los valores temáticos o atributos de la consulta.

En el caso de los sistemas ráster este tipo de consultas nos muestran todos los píxeles que se encuentran en el recinto de la consulta, a los que podríamos asociar los valores temáticos de otros mapas superpuestos en esas posiciones.

2.2. Consultas de Relaciones Espaciales

Estas consultas nos permiten seleccionar entidades basándonos en la posición que éstas ocupan con respecto a otra u otras.

La expresión de la consulta de forma genérica podría ser[140]:

Selecciona *expresión sujeto* operador espacial *expresión predicado*

En esta sintaxis la palabra *expresión* hace referencia a cualquier combinación formada por clases de objetos combinados con operadores lógicos y/o resultado de otras consultas. El *sujeto* y el *predicado* son clases de objetos.

Un operador espacial es la expresión formalizada de las condiciones que han de cumplir dos entidades o conjuntos de entidades para encontrarse en una cierta situación espacial relativa[141]. Nos permiten analizar las situaciones espaciales entre objetos, y verificar si se cumplen las condiciones de la consulta que hacen referencia a una determinada situación o interacción espacial entre objetos. Los operadores se definen por medio de algoritmos que son los que calculan si se cumplen las condiciones de un operador haciendo uso de la posición o geometría de los objetos, o lo que es lo mismo, las primitivas topológicas o geométricas.

Algunos operadores espaciales son: intersecta, cruza, toca, contiene, dentro, solapa, igual, o disjunto. Los operadores también se pueden crear por combinación de los anteriores.

Por lo general las aplicaciones SIG actuales ayudan a realizar este tipo de consultas por medio de una interfaz en la que se muestran las

140. Pilar Moreno Regidor, 'Fundamentos de Sistemas de Información Geográfica', transparencias adaptadas al programa de la asignatura de la E. U. I. T. Topográfica. E. T. S. de Ingeniería en Topografía, Geodesia y Cartografía. UPM, Madrid., 2006.
141. Francisco J. García Lázaro, 'Sistemas de Información Geográfica', transparencias adaptadas al programa de la asignatura de la E. U. I. T. Topográfica. E. T. S. de Ingeniería en Topografía, Geodesia y Cartografía. UPM, Madrid., 2001.

combinaciones de las relaciones espaciales que se pueden utilizar en este tipo de consultas como se muestra en la Figura 11.

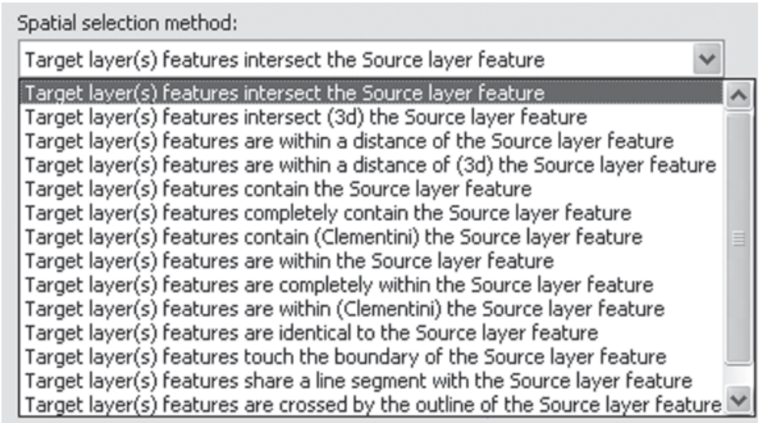


Figura 11. Reglas de las relaciones espaciales entre clases de objetos.
Fuente: ESRI[142].

En muchas ocasiones en este tipo de consultas se utilizan criterios temáticos para seleccionar los conjuntos de objetos sobre los que se va a realizar la selección por relación espacial, por lo que en realidad se deberían considerar como *consultas mixtas*.

142. ESRI, 'Seleccionar por ubicación', 2010. <http://help.arcgis.com/es/arcgisdesktop/10.0/help/index.html#//00s50000002q000000>

Capítulo V
Análisis espacial

Este capítulo aborda el tema del análisis espacial que, junto con las consultas —desarrolladas en el capítulo anterior—, son las dos piezas clave de cualquier SIG, donde reside su mayor potencialidad. Hay que diferenciar entre lo que son funciones de análisis propiamente dichas y aquellas destinadas a la manipulación o transformación de los datos previas a dicho análisis, y que no se tratarán en este capítulo; tal es el caso de operaciones como la generalización cartográfica, el ajuste de hojas, los cambios en los sistemas de proyección, etc.

Dada la extensión del tema que nos ocupa —inabordable en un libro de estas características— se van a enumerar y explicar brevemente tan solo las principales, tratando de ejemplificar su aplicación con algunos casos concretos, con el objetivo de aportar una visión general de las enormes capacidades analíticas de esta tecnología especialmente en el campo de las Ciencias Humanas y Sociales.

Se entiende como *análisis espacial*[143] el conjunto de procedimientos de estudio de los datos geográficos en los que se consideran de alguna manera sus características espaciales; es por tanto una colección de procesos cuyo fin es explotar nuestros datos geográficos.

La forma de abordar el análisis espacial difiere en el modelo vectorial y ráster por la propia naturaleza de los datos en cada caso. En líneas generales, hay más funcionalidades de análisis para los datos rás-

143. David J Unwin, *Introductory Spatial Analysis.* London: Menthuen, 1981.

ter que vectoriales porque la aplicación de algoritmos es más sencilla en éstos gracias a su estructura en píxeles.

A continuación se tratarán por separado el análisis espacial en el mundo vectorial y ráster; si bien en ambos casos se desarrollarán conceptos muy similares (mapas de distancias, superposición de capas,…) la forma de resolverlos como veremos será completamente diferente.

Cerraremos el capítulo dedicando algunos apartados a otras cuestiones relacionadas con el análisis espacial pero no tan vinculadas al modelo de representación en sí, como son el concepto de autocorrelación espacial, los algoritmos de interpolación, las técnicas de evaluación multicriterio o lo que se conoce como análisis exploratorio espacial de datos, entre otros.

Análisis Vectorial

Son muchas las funcionalidades de análisis SIG vectorial, unas se centran en el estudio descriptivo de las variables temáticas, otras en el análisis puramente espacial, y por último, las más complejas analizan a la vez tanto la componente geográfica como la temática, otorgando al SIG una elevada potencialidad analítica.

En un SIG, en el caso de necesitar analizar exclusivamente las variables temáticas sin tener en cuenta su relación con el espacio o con otros objetos, se puede abordar como un análisis estadístico clásico, aunque sin perder de vista las particularidades de los datos espaciales (existencia de autocorrelación espacial, problema de la unidad espacial modificable, etc.)[144].

Si lo que se pretende por el contrario es centrar el análisis solamente en la componente espacial, las funcionalidades se centran en analizar por ejemplo las características geométricas de las entidades geográficas objeto de estudio (longitud de líneas, perímetros o áreas de polígonos, forma del polígono mediante el índice de circularidad…) o en realizar cálculos de distancias y análisis de proximidad.

144. Joaquín Bosque Sendra, *op. cit.*, p. (127).

No obstante, lo más habitual es querer analizar a la vez ambas componentes, espacial y temática. A continuación se desarrollan algunos bloques temáticos de importancia en la materia que nos ocupa:

- Análisis descriptivo de la componente temática.
- Análisis de mapas de puntos.
- Superposición de mapas vectoriales.
- Cálculo de distancias y análisis de proximidad.
- Análisis de redes.

Análisis descriptivo de la componente temática.

Los elementos de la estadística clásica tienen sus equivalentes en los datos espaciales (que como ya hemos visto tienen características muy particulares), y nos permiten calificar cuantitativamente los datos con los que trabajamos. Se incluyen aquí descriptores de centralidad (como el centro medio, el centro de gravedad) y dispersión (como la distancia típica o el radio dinámico), de dependencia espacial o el estudio de patrones espaciales, entre otros muchos. Estos pueden a su vez usarse para el contraste de hipótesis que contengan una cierta componente espacial[145].

Algunas de las cuestiones tipo que responderían estos estadísticos son: ¿los individuos pertenecientes a un determinado gremio en el pasado tenían tendencia a agruparse o por el contrario se dispersaban por todo el territorio?, ¿existe alguna dirección predominante en los movimientos de individuos de una especie o se desplazan erráticamente?, ¿están dispuestos del mismo modo en el espacio los centros de educación públicos que los privados?, etc.

En el caso de las variables cuantitativas, tan utilizadas por ejemplo en las ciencias sociales (censos, datos de encuestas, etc.), es muy habitual para entender un fenómeno geográfico determinar la relación entre dos variables, analizando también el sentido y la intensidad de la relación si existiera. Para ello se realizan *análisis de regresión* y se

145. Víctor Olaya, *op. cit.*

aplican estadísticos como el coeficiente de correlación de Pearson (r) que mide la fuerza de la relación en caso de existir; ésta se visualiza con una línea de regresión donde la pendiente mide la velocidad del cambio. Aunque dedicaremos un bloque a esto, adelantar que se puede hablar de *correlación positiva* cuando los valores de ambas variables aumentan, *negativa* cuando los valores de una aumentan y los de la otra disminuyen o *nula* cuando no se relacionan.

Otro ejemplo de aplicación, esta vez en el campo de la Arqueología tiene que ver con la utilización de análisis de regresión de variables medioambientales (altitud, distancia a fuentes de agua) en relación con la localización de yacimientos, o el análisis entre el porcentaje de un material como la cerámica (variable dependiente) en cada yacimiento y la distancia de éste al centro de producción (variable independiente)[146].

Análisis de mapas de puntos.

Es muy interesante en algunos estudios conocer la distribución de una serie de puntos en el espacio (ej.: aparición de patrones de distribución de una enfermedad y del factor causal, ubicación de individuos, presencia o ausencia de una determinada variable en el territorio, distribución espacial de yacimientos o asentamientos para analizar la posible competencia entre ellos, etc.) en tanto en cuanto aportan información valiosa acerca de las variables objeto de la investigación.

En líneas generales, una distribución de puntos puede ser de tipo aleatoria, regular o agregada. Los estadísticos buscan caracterizar la disposición observada de un conjunto de puntos como la más próxima a una de esas tres posibilidades teóricas. No hay que olvidar que en este tipo de análisis es muy importante el tamaño del área de estudio para detectar patrones a escalas diferentes.

Algunos de los análisis estadísticos que se realizan son el análisis de cuadrantes aplicando tests como el de la Ji^2 o el de Kolmogorov-Smirnov (dividiendo la zona de estudio en unidades regulares y estu-

146. Ian Hodder y Clive Orton, ***Análisis espacial en Arqueología.*** Barcelona: Crítica, 1990.

diando el número de puntos que aparecen dentro de cada una), el método de vecino más próximo (que se basa en las distancias de cada punto a su vecino más cercano, de forma que comparando estas distancias con el valor que cabe esperar en una distribución aleatoria, puede deducirse el tipo de estructura en la distribución observada) o la función K de Ripley (que trata de incorporar la escala como elemento de análisis), entre otros[147].

Superposición de mapas vectoriales.

En un modelo vectorial la superposición de capas implica resolver tanto la parte geométrica como la temática en la capa de salida. Como vimos en el apartado de las consultas espaciales, las combinaciones posibles entre mapas son muy variadas (puntos con polígonos, líneas con polígonos, líneas con líneas, polígonos con polígonos, etc.).

El caso más complejo a nivel geométrico es quizás el de la superposición entre capas poligonales. Se realiza en dos fases: una geométrica o topológica, donde se generan nuevos polígonos, se les asigna nuevos identificadores y se rehace la topología de la capa, y una temática, en la que se asignan valores a los atributos en función de los de las capas de entrada. Esta última fase se resuelve de diferente forma si las variables de entrada son nominales, ordinales o cuantitativas (mediante álgebra de mapas, enmascarado de mapas, interpolación areal, etc.).

En la superposición de capas poligonales un problema frecuente que se ha de tratar de evitar es la aparición de polígonos ficticios (*slivers*), que son producto de errores de digitalización al intersectar capas de, en teoría, límites poligonales iguales. Si aparecen hay que eliminarlos al resolver la fase topológica, estableciendo un margen de tolerancia, ya que suelen ser de tamaño muy reducido.

Cálculo de distancias y análisis de proximidad.

El concepto de *distancia* entre dos puntos o entidades se utiliza en multitud de análisis espaciales. En el caso de los modelos vectoriales

147. Víctor Olaya, *op. cit.*

ésta se expresa como una función matemática y se definen dos tipos: la distancia *euclídea* (que sigue el teorema de Pitágoras midiendo la longitud de la línea recta que une dos puntos) o la distancia de *Manhattan* (más apropiada en un tramado urbano donde el movimiento ha de seguir el trazado de las calles). En el caso de necesitar realizar el cálculo de distancia entre dos puntos alejados de la superficie terrestre, la Tierra no puede considerarse plana, por lo que habría que tener en cuenta la distancia geodésica.

En ocasiones, más que de distancia se habla de *coste de recorrido* entre un punto A y otro B; en esta caso no siempre es la longitud la que indica dicho coste, a veces se debe a otros factores (tiempo, gasto económico, esfuerzo físico, etc.).

Derivado del concepto de la distancia existen dos funcionalidades muy habituales en cualquier SIG: la generación de áreas de influencia (*buffer*) o análisis de vecindad, y la creación de polígonos de Voronoi o Thiessen.

Para finalizar con los análisis relacionados con el concepto de distancia, está la generación de *áreas de influencia* (*buffer*), también conocido como análisis de vecindad o de proximidad. Es un tipo de consulta espacial en el que se define un área alrededor de un objeto geográfico puntual, lineal o poligonal, especificando la anchura de la misma, o varias en el caso de querer establecer corredores de influencia concéntricos. Es un tipo de análisis más complejo en el modelo vectorial que en el ráster porque el sistema ha de dibujar los polígonos, generar topologías, resolver en su caso las intersecciones, etc.[148]

Este tipo de análisis responde a cuestiones del tipo: ¿qué porcentaje de yacimientos se halla a menos de 2 Km de distancia de la línea de costa?, ¿qué instalaciones recaen en el municipio X y además están a menos de 50 m de una carretera?

148. Javier Gutiérrez Puebla y Michael Gould, *op. cit.,* p. 177.

Creación de polígonos de Voronoi o de Thiessen.

Es un tipo de análisis en el que se crea una capa poligonal a partir de una capa de puntos basándose en el concepto de la distancia mínima entre puntos, de forma que los lados de éstos son equidistantes con respecto a los puntos vecinos. Son útiles para generar mapas de coropletas a partir de capas puntuales (ej.: mapas de suelos en base a los resultados obtenidos de catas puntuales in situ).

Así, una distribución de puntos se convierte en un conjunto de elementos poligonales o áreas de influencia de cada uno de ellos donde, cada punto dentro del mismo, pertenece a éste por encontrarse a la menor distancia (también se pueden incluir valores de ponderación de los puntos en base a un atributo ej.: la población).

Los polígonos Voronoi se han utilizado por ejemplo en demografía histórica para crear los ejes de calles en el proceso de vectorización de un plano de Madrid de 1902[149].

Análisis de redes.

Una *red* es un sistema de elementos lineales interconectados por los que pueden circular flujos de diverso tipo, tangibles (personas, mercancías, objetos antiguos, energía) o intangibles (información)[150].

Las funcionalidades más usuales en el análisis de redes se refieren al cálculo de caminos mínimos (se conoce como *routing*) –menor impedancia implica menor resistencia al movimiento- y al análisis de áreas de influencia de centros de servicio o actividades (conocido como *allocate*), que puede realizarse como hemos visto en el apartado anterior utilizando polígonos Voronoi o mejor a través de una geometría de red real[151].

149. Isabel Del Bosque González et al., *op. cit*. Este procedimiento aparece anteriormente descrito por Emilio Gómez Fernández (2008). y se puede consultar en la web: http://es.wikipedia.org:80/wiki/Archivo:Voronoi_centerlines_skeleton.gif. [accedido 13 Junio 2011].

150. Joaquín Bosque Sendra, *op. cit.,* p. (207).

151. Javier Gutiérrez Puebla y Michael Gould, *op. cit.*, p. (185).

Otro concepto es el de las redes de flujo, cuando lo que interesa es representar el movimiento de bienes entre dos lugares (dirección y volumen). Entonces se introducen ponderaciones como es el caso de las impedancias (representan limitaciones al movimiento como p.ej.: el tipo de asfalto en una red de caminos) o las máximas capacidades de circulación de un determinado bien[152].

En investigación histórica se está empezando a utilizar este tipo de análisis para por ejemplo reconstruir rutas mercantiles de la antigüedad a partir de la distribución de hallazgos de bienes importados y procedentes de depósitos, o para analizar la creación y mantenimiento de complejas redes sociales comerciales, o para estudiar la influencia de la proximidad física en la formación de algunas de estas redes, entre otros. Este tipo de investigaciones son de gran complejidad por el elevado número y diversidad de factores que intervienen en la formación de la red. Como comprobaremos posteriormente, el modelo vectorial es más útil en este tipo de estudios que el ráster.

También es muy común utilizar análisis de redes vectoriales en el caso de los estudios de localización-asignación, que persiguen asignar la localización óptima de una instalación (hospital, escuela, sucursal bancaria, etc.), conociendo tanto su distribución espacial como los elementos que influyen en su utilización.

Análisis Ráster

Como ya se adelantó en el capítulo de fundamentos, en el caso del modelo ráster no existe una separación tan clara entre las componentes espacial y temática como ocurre en el caso del modelo vectorial, lo que implicará por ejemplo que a la hora de superponer capas no haya que resolver la parte geométrica porque por norma general esta es única y coincidente en todas ellas (el píxel es la unidad espacial mínima), facilitando por tanto este tipo de operaciones.

El análisis se puede realizar bien sobre un solo mapa o bien sobre varios a la vez superponiéndolos. A la hora de realizar análisis con rás-

152. James Conolly y Mark Lake, *op. cit.*

ter existen básicamente dos aproximaciones: tener en cuenta la totalidad del mapa —se conoce como *análisis global*—, o centrar el análisis tan solo en algunas localizaciones específicas que podrían ser píxeles (en las operaciones píxel a píxel no se tienen en cuenta las relaciones espaciales de ese con respecto a los otros), vecindades ó zonas[153] —se denomina *análisis local*—. Por último las funciones de análisis ráster pueden operar con los atributos temáticos (ej.: reclasificaciones, obtención de estadísticos, etc.) o con los aspectos espaciales[154] (ej.: forma y tamaño, distancia y proximidad, vecindades, etc.).

Son muchas y muy diversas las clasificaciones de las funcionalidades de análisis ráster de un SIG. Una de ellas las divide en cuatro grandes grupos[155] que se desarrollan a continuación:

- Reclasificación de los valores temáticos de un mapa.
- Superposición de mapas.
- Cálculo de distancias y de conectividad.
- Caracterización en base a las vecindades.

Reclasificación de los valores temáticos de un mapa.

Es una operación que se realiza sobre un único mapa y consiste en la generación de un nuevo mapa modificando los valores de la variable temática; no existe por tanto creación de nuevos límites para las entidades. La asignación de esos nuevos valores puede realizarse píxel a píxel o teniendo en cuenta los valores de vecindades o zonas.

Algunos ejemplos de reclasificaciones son:

- los *reetiquetados de valores* (ej.: si en un mapa de usos del suelo la categoría de urbano está codificada con el valor "5" y deseamos cambiarla a valor "4"),

153. Se entiende por *vecindad* el conjunto de píxeles que rodean uno de referencia, a una distancia o en una dirección concreta; y por *zona* o *región* el conjunto de píxeles con idéntico atributo temático, independientemente de si son o no contiguos.

154. Joaquín Bosque Sendra, *op. cit.*, p. (310).

155. Joseph Berry, 'Fundamental operations in computer-assisted map analysis', *International Journal of Geographical Information Science*, 1 (1987) pp. (119-136).

- las *agregaciones de categorías* iniciales en un número menor (ej.: si los píxeles etiquetados inicialmente como “1=olivar”, “2=frutal”, y “3=cereal”, se agregan en una única categoría codificada como “1=cultivo”),
- las *creaciones de intervalos* en el caso de variables continuas pasando de variables cuantitativas a cualitativas (ej.: reclasificar un modelo digital de elevaciones, que en cada píxel guarda la altitud del terreno en ese punto, en intervalos del tipo “1 = nivel del mar a 500 m”, “2 = 500 a 1000 m”, “3 = 1000 a 1500 m”, “4 = más de 1500 m”),
- las *operaciones con una constante* utilizando operadores como la suma, resta, producto, cociente, etc. (ej.: para pasar un mapa de distancias que está medido en kilómetros a otro medido en metros debemos multiplicar el valor de cada píxel por 1000),
- el *análisis de parámetros geométricos de las entidades* tales como la forma, la superficie, la convexidad, etc. (ej.: en un estudio urbano puede interesar reclasificar los píxeles de suelo urbanizado asignándoles un coeficiente que haga referencia a su forma, como es el caso de la razón de circularidad, que mide la forma geométrica de una zona o región).
- el *enmascarado de mapas* (uno de los mapas actúa de máscara sobre el otro, de forma que el mapa final toma los valores de la máscara salvo en el caso de los píxeles con un valor seleccionado como el nulo, donde toma los del segundo mapa).

Superposición de mapas (overlay).

Es una operación que se realiza sobre varios mapas a la vez y, al igual que en el caso de las reclasificaciones el análisis puede ser local, regional o zonal. Las superposiciones entre mapas se pueden realizar bien utilizando operadores lógicos (AND, OR,…) o bien utilizando operaciones aritméticas (suma, resta, cociente,…). El resultado final es un nuevo mapa donde los valores de los píxeles dependen de los valores de éstos en cada una de las capas que se superponen.

Otro aspecto diferenciador es si se manejan variables nominales o cuantitativas. En el caso de las primeras el proceso se asemeja al de la

tabulación cruzada, donde se genera una tabla que resume el número de píxeles que pasa de una categoría a otra y el número de píxeles que no ha sufrido cambio alguno.

Un ejemplo de aplicación podría ser el del análisis de paisajes culturales o históricos, que trata de identificar y analizar la estructura y el cambio del uso del suelo en el tiempo. El análisis de estos cambios podría marcar por ejemplo, en investigaciones arqueológicas, los lugares donde la reducción de usos tradicionales implicó la pérdida de hábitats para la flora y la fauna propia de esas zonas[156].

En el campo del análisis socioespacial, se ha utilizado la superposición de capas ráster por ejemplo para analizar y cuantificar la pérdida de suelos productivos agropecuarios a favor del crecimiento urbano, cruzando para ello mapas de superficie urbana a lo largo de varios años, con el fin de obtener mapas de crecimiento absoluto y relativo de la aglomeración y pérdida de suelo de mayor productividad agropecuaria[157].

Cálculo de distancias y de conectividad.

Es una operación que consiste básicamente en medir distancias entre puntos. Al igual que lo que vimos en el caso del modelo vectorial, en un mapa ráster se pueden medir tanto la distancia euclídea como la de Manhattan (calculando el número de píxeles que separa un punto de otro).

Y, como ya vimos, la distancia entre dos puntos no siempre se mide en unidades de longitud, sino que tiene que ver con el "coste" de recorrer esa longitud que los separa (que se puede medir en términos de tiempo, esfuerzo físico, gasto monetario, consumo energético, etc.). Un *mapa de costes* guarda por tanto, para cada píxel, el valor de atravesarlo en términos de coste, y permite definir ciertas zonas como barreras ab-

156. Oliver Bender, 'The workflow of a historic landscape analysis using GIS with examples from Central Europe', en *Geoinformation Technologies for Geocultural Landscapes. European Perspectives*London, UK: CRC Press, 2009, pp. (171-187).

157. Gustavo D. Buzai y Claudia A. Baxendale, *op. cit.*

solutas (a las que se les asigna un coste altísimo para impedir que sean atravesadas, p.ej.: una cadena montañosa o un río) o barreras relativas (a las que se les asigna un coste elevado aunque permitiendo ser atravesadas, p.ej.: una carretera cuyo firme está en mal estado).

Los mapas de costes de recorrido se clasifican en isotrópicos (dependen de la dirección del viaje), parcialmente anisotrópicos (dependen de la dirección del viaje, pero la dirección de máximo coste es la misma para todas las celdas del mapa), o anisotrópicos (dependen tanto de la dirección del viaje como de los atributos de cada una de las celdas del mapa)[158].

En estudios de Arqueología del Paisaje, por ejemplo, es crucial conocer la distancia entre dos asentamientos a la hora de analizar la relación entre ellos o la de éstos con el resto; para ello se realizan mapas de coste. A modo de ejemplo, en una investigación que se realizó para estudiar la organización social y económica y las dinámicas de las sociedades preindustriales en entornos mediterráneos, se valoró el coste en función de la pendiente y dichos mapas se combinaron con otros de cuencas visuales[159].

Otra aplicación derivada del cálculo de distancias en mapas ráster es el cálculo del *camino óptimo entre dos puntos*; en este caso se trata de disminuir el coste de recorrido en la mínima distancia que separa dos puntos. Como ya vimos anteriormente, este tipo de cálculos es más sencillo en el modelo vectorial utilizando redes.

Por último, y relacionado también con el concepto de conectividad entre objetos, los *análisis de intervisibilidad* o el *cálculo de cuencas visuales* son herramientas analíticas muy potentes. En este caso se precisa un modelo digital de elevaciones y la posición del objeto/s que hace de observador/es; el SIG calculará tanto las zonas visibles y ocultas desde cada punto, como las zonas que ven o no a un observador

158. James Conolly y Mark Lake, *op. cit.*
159. Teresa Chapa Brunet et al., 'GIS landscape models for the study of preindustrial settlement patterns in Mediterranean areas', en *Geoinformation Technologies for Geocultural Landscapes. European Perspectives.*London, UK: CRC Press, pp. (255-273).

concreto. Una posible aplicación son los estudios de impacto visual de nuevas instalaciones, o el análisis del campo visual desde un yacimiento arqueológico concreto.

Caracterización en base a las vecindades.

Este epígrafe se refiere al conjunto de operaciones que generan un nuevo mapa donde el valor asignado en una posición es el resultado de aplicar una función que tiene en cuenta los valores individuales de los píxeles que le rodean, es decir, de su vecindad; siendo el tamaño y forma de ésta variable en cada caso.

Una aplicación de esto son los *filtrados de mapas*. Con los filtros espaciales se calcula el valor de una celda en función de sus vecinas y el resultado en ocasiones es algún tipo de promedio ponderado. Por lo general se usa una vecindad en forma de ventana cuadrada que se desplaza por el mapa asignando a cada píxel un nuevo valor filtrado. En teledetección es muy común aplicar filtros para realzar las imágenes o un determinado rasgo de las mismas (filtros de paso alto), o para suavizarlas (filtros de paso bajo).

Otro ejemplo son las funciones que tienen como objeto resumir una determinada propiedad en base a los valores de la vecindad, tal es el caso de los análisis de diversidad (que devuelven, para cada píxel, el número de variantes de un determinado atributo –ej.: categorías de uso de suelo- que caen dentro de la ventana que define la vecindad), la extracción de la categoría más o menos común, el cálculo del valor medio, la varianza o la desviación típica de un conjunto de valores, etc.

Para finalizar el capítulo, abordaremos otras cuestiones como son el concepto de autocorrelación espacial, los algoritmos de interpolación, las técnicas de evaluación multicriterio o lo que se conoce como análisis exploratorio espacial de datos.

Autocorrelación Espacial

La primera ley de la geografía, o *principio de autocorrelación espacial*, fue formulado en 1970 por el geógrafo Waldo Tobler de esta forma: "*Todas las cosas están relacionadas entre sí, pero las cosas más próximas en el espacio tienen una relación mayor que las distantes*"[160], es decir, la presencia de un valor concreto de una variable hace que ese valor se repita o aproxime en localizaciones próximas. Es un hecho que se da con más frecuencia en variables físico-naturales que socioespaciales.

Su estudio busca responder a cuestiones del tipo: ¿cómo influye la componente espacial en la distribución espacial de una variable?, ¿la distribución de una característica socioespacial se produce de forma aleatoria o existe cierta autocorrelación?, o ¿qué ocurre si se analiza su distribución en relación con otra característica (bivariante) u otras (multivariante)?.

Se habla de *autocorrelación positiva* si el valor en las unidades espaciales próximas es similar (se producen fenómenos de agrupamiento), de *autocorrelación negativa* en caso contrario (se produce dispersión), o puede ser que simplemente no exista.

Existen diversos estadísticos que analizan si hay autocorrelación positiva o negativa, así como otros que definen la forma de dicha autocorrelación, como es el caso del variograma, o los que la cuantifican como el estadístico I de Moran, que expresa de manera formal el grado de asociación lineal entre dos variables aleatorias independientes, comparando los valores de cada localización con los valores de las localizaciones contiguas.

Un ejemplo de aplicación en Arqueología sería averiguar si el valor de la densidad de artefactos en una zona es aleatoria o está relacionada con la presencia de ciertos rasgos; a priori a lo largo de ríos y caminos se produce un aumento del movimiento y la interacción, que afectan a la distribución de artefactos[161].

160. Waldo R.Tobler, 'A Computer Movie Simulating Urban Growth in the Detroit Region', *Economic Geography*, 46 (1970) pp. (234-240).

161. James Conolly y Mark Lake, *op. cit.*

Interpolación Espacial

El proceso de *interpolación espacial* consiste en la estimación de los valores que alcanza una variable Z en un conjunto de puntos definidos por un par de coordenadas (X,Y), partiendo de los valores de Z medidos en una muestra de puntos situados en el mismo área de estudio (por ejemplo la interpolación de la temperatura a partir de datos de observatorios meteorológicos)[162]. Son muchos los métodos de interpolación, pero todos ellos asumen la existencia de autocorrelación espacial positiva; requisito imprescindible para poder interpolar.

En un estudio espacial sólo disponemos de los valores medidos en algunos puntos de muestreo en un espacio infinito, por tanto la estimación de medias, desviaciones típicas y covarianzas no es muy fiable. Es por esto que se utiliza otro estadístico alternativo al coeficiente de correlación, el semivariograma, que da una visión de cuál es la estructura de variabilidad espacial de una variable medida en un conjunto de puntos.

Una primera clasificación de las *funciones de interpolación* las divide en[163]:

- *Globales*: tienen en cuenta todos los valores muestrales para identificar tendencias globales (ej.: análisis de superficies de tendencia).
- *Locales*: solo tienen en cuenta los valores conocidos de una vecindad determinada (ej.: interpolación lineal con ponderación del inverso de la distancia IDW, útil para datos muy espaciados, o los *splines*, indicados para superficies que varían leve y lentamente como la altitud, o el *kriging*, apropiado para modelizar con precisión datos que tengan un comportamiento uniforme en toda la zona considerada, e inapropiado para modelizar fenómenos que tengan rupturas importantes o grandes cambios abruptos, o los polígonos Voronoi, en el caso de variables cualitativas).

162. A. Stewart Fotheringham, Chris Brunsdon y Martin Charlton, *Quantitative Geography: Perspectives on Spatial Data Analysis.* London: Sage Publications, 2000.

163. Lubos Mitas y Helena Mitasova, 'Spatial interpolation', en *Geographical Information Systems: Principles, techniques, management and applications.* New Jersey: John Wiley & sons, Inc., 2005, pp. (481-492).

La elección y aplicación de una u otra depende de varios factores (naturaleza de los datos, propósito de la interpolación) y los resultados van a variar.

Una de las aplicaciones inmediatas es la generación de modelos digitales de elevaciones a partir de curvas de nivel. Los MDE son muy importantes por ejemplo en investigación geoarqueológica y geocultural porque aportan información sobre algunos parámetros morfométricos de gran importancia (pendientes, orientaciones, curvaturas) que ayudan a entender la estructura del paisaje y sus procesos.

Técnicas de Evaluación Multicriterio (EMC)

La *evaluación multicriterio* (y multiobjetivo) es un conjunto de técnicas utilizadas en la decisión multidimensional y los modelos de evaluación, dentro del campo de la toma de decisiones[164]. Consiste en la evaluación de alternativas para optar por la que mejor satisface las condiciones establecidas; todas han de quedar evaluadas y ordenadas según su grado de adecuación. Es una herramienta muy potente en la toma de decisiones en el campo por ejemplo de la planificación.

Los elementos que componen cualquier EMC son: la *decisión* (alternativa elegida), las *alternativas* (soluciones posibles al problema planteado) y los *criterios* (condicionantes que deben cumplir las soluciones). Estos últimos pueden ser factores (cuantitativos u ordinales) o restricciones (siempre binarias, sí/no), y se les puede ponderar asignándoles pesos en función de la importancia que les da la persona que tomará las decisiones. Los criterios se combinan según las llamadas reglas de decisión.

Los criterios se modelan en forma de capas geográficas (*layers*) como apoyo a la toma de decisiones. Para facilitar su manipulación, en ocasiones las capas o criterios se transforman en capas booleanas actuando como factores, todas con el mismo peso, que se combinan después mediante una intersección booleana utilizando la multiplica-

164. José Ignacio Barredo Cano, *Sistemas de información geográfica y evaluación multicriterio: en la ordenación del territorio.* Madrid: Ra-Ma Editorial, 1996.

ción). Otra forma de combinarlas es hacer capas de factores según lo que se denomina *lógica difusa* (*fuzzy*), en donde cada píxel se clasifica en diferentes niveles entre “apto” y “no apto”[165].

La EMC se aplica por ejemplo en los denominados *modelos de localización-asignación* (búsqueda de la localización óptima de un determinado servicio o instalación y cálculo de la asignación de demanda a la misma). En estudios de análisis socioespaciales, un ejemplo de aplicación consistiría en la búsqueda de los sitios de mayor aptitud para la instalación de por ejemplo una industria, o un centro de servicio demandado por la población (hospital, bomberos, escuela…) en estudios de mercado que analizan la competencia o la oferta frente a la demanda.

Análisis Exploratorio Espacial de Datos (ESDA)

El análisis exploratorio de datos espaciales (ESDA)[166] es un conjunto de técnicas destinadas a detectar esquemas de asociación espacial, concentraciones locales y regímenes espaciales presentes en un conjunto de datos para los que las características de localización resultan esenciales[167]. Sirve por tanto para describir y visualizar las distribuciones espaciales de esos datos e identificar los esquemas de asociación espacial dominantes, agrupamientos (*clusters*), las localizaciones atípicas (*spatial outliers*), las tendencias globales de los datos, o la heterogeneidad espacial, entre otros.

El ESDA combina la estadística y los SIG, busca conocer en una primera aproximación el comportamiento de una variable, o la relación entre varias (utilizando medidas de centralidad, diagramas de dispersión 2D y 3D combinados con la cartografía, etc.) en su contexto espacial, aunando por tanto la componente espacial y temática de los datos en el análisis.

---

165. Gustavo D. Buzai y Claudia A. Baxendale, *op. cit.*

166. El ESDA es una subdisciplina del Análisis Exploratorio de Datos diseñada para el tratamiento específico de los datos geográficos por las particularidades que presentan.

167. Luc Anselin, ‘Exploratory spatial data analysis and geographic information systems’, en *New tools for spatial analysis*Luxembourg: Eurostat, 1994, liv, 45-54 <http://www.rri.wvu.edu/pdffiles/wp9329.pdf> [accedido 15 Agosto 2011].

En el campo de las Ciencias Sociales se utiliza con frecuencia en modelos econométricos. Algunas de las herramientas estadísticas que se utilizan son los gráficos de caja, los diagramas de dispersión de Moran (uni y multivariantes) para ver el tipo de asociación espacial, o los mapas de histogramas de frecuencias para analizar la heterogeneidad espacial.

Modelos Cartográficos

Ante proyectos SIG de gran complejidad, donde la componente analítica cobra mucho peso, es muy útil el empleo de herramientas que nos ayuden a definir modelos cartográficos mediante una notación gráfica, que detallan los datos de entrada, los de salida, y las operaciones de transformación o analíticas intermedias.

Una de las ventajas de estos modelos es la posibilidad de simular diversos escenarios ejecutando el modelo de forma repetitiva modificando ciertos parámetros o variables de entrada. Otra ventaja es la flexibilidad, permite añadir nuevas consideraciones al modelo o hacer pequeñas modificaciones en las existentes. También es una herramienta de comunicación con el resto de integrantes de un proyecto (la lógica, las relaciones, las variables de entrada, las condiciones, etc.), favoreciendo la participación de las personas en el proceso analítico.

La implementación del modelo en muchos SIG se realiza de forma automática (ej.: módulo ModelBuilder en ArcGIS®).

Capítulo VI
Representación de los datos

A diferencia de los mapas en papel, los SIG separan el almacenaje de los datos de su representación. Es posible por lo tanto obtener variadas representaciones sobre un mismo conjunto de datos, en función del mensaje que queramos transmitir mediante el mapa, un proceso tradicionalmente muy costoso, facilitado hoy día por la accesibilidad de tecnología que permite a cualquier usuario de un SIG generar una salida gráfica a la escala de su elección y con la simbología que desee en cuestión de minutos.

Las herramientas SIG no sólo permiten al usuario final de un mapa su elaboración personalizada, también abren un gran abanico de posibilidades, desde la creación de un tradicional mapa topográfico o temático hasta la generación de animaciones o mapas dinámicos que muestren la evolución de una variable a lo largo del tiempo.

La importancia de la Generalización

Antes de proceder a generar una representación, para que ésta sea clara y legible debemos cuidar la estructura de la información que vamos a mostrar. La representación implica un grado de reducción de información establecido por la escala, que limita claramente el espacio disponible sobre el mapa.

La selección de información relevante de una base de datos geográfica implica una abstracción que depende del entendimiento de los conceptos geográficos. Es lo que se llama *generalización semántica*.

Se preocupa del significado y función del mapa, y depende de ser capaz de identificar estructuras jerárquicas en la información geográfica; una base de datos bien organizada nos permitirá elegir los elementos gráficos y textuales que comuniquen adecuadamente el significado de los datos. La parte de la *generalización geométrica*, consistente en aumentar el nivel de abstracción gráfica en relación a la forma original de los datos ya fue expuesta en el capítulo de Edición y Consulta. Las estructuras implicadas en la generalización semántica pueden ser de clasificación o agregación[168].

Las clasificaciones se aplican a muchos fenómenos naturales como la hidrografía, por ejemplo, que puede ser subdividida en lagos, océanos, ríos, etc. Los esquemas de clasificación también pueden ser cuantitativos, basados en intervalos o escalas de medida. Un claro ejemplo sería una clasificación de ciudades en intervalos según su población, para establecer una representación diferenciada o excluir de nuestra representación alguna categoría.

Los esquemas de agregación se usan para representar objetos en términos de sus partes constitutivas, que pueden ser subdivididas. Por ejemplo, en una ciudad se contemplarían, por un lado los distritos, la red de calles, los bloques domésticos, comerciales, industriales, de ocio... existen todas las partes por separado y se pueden agrupar en función de niveles de interés.

Los mapas temáticos en Ciencias Humanas y Sociales

En el terreno de las Humanidades y las Ciencias Sociales los mapas temáticos constituyen un apoyo básico para la contextualización geográfica de los fenómenos estudiados. Aunque el nivel de implicación con las tecnologías de información geográfica varía en función de la disciplina[169], la representación de datos como un primer paso en el análisis o como visualización de resultados es habitual en la mayoría

168. Christopher Jones, *Geographical Information Systems and Computer Cartography.* Essex: Longman, 1997.
169. *Geospatial Technology and the Role of Location in Science*, ed por Henk J Scholten, Rob Velde y Niels Manen, 1o edn. Dordrecht: Springer Netherlands, 2009.

de campos, desde la Lingüística hasta la Arqueología, pasando por la Demografía, la Historia, las Ciencias Económicas o, por supuesto, la Geografía.

En los mapas temáticos la componente topográfica se simplifica, convirtiéndose en lo que se denomina un "mapa base", que ofrece la contextualización geográfica necesaria para entender las variables representadas: movimientos migratorios, información histórica, agrícola y forestal, político administrativa, etc.

Los mapas temáticos pueden representar variables cuantitativas o cualitativas y basar su representación en puntos, líneas o polígonos, ubicados generalmente sobre un mapa base[170].

- *Mapas cualitativos:* son mapas que representan condiciones o características de los elementos representados. Cuando se trata de mapas de puntos se utilizan símbolos geométricos o pictóricos, en los casos de fenómenos lineales o superficiales se utilizan tramas o colores para mostrar las diferencias en las cualidades de la información representada.

  Son muy utilizados para representar información puntual en el caso por ejemplo de elementos clasificados en función de algún atributo; ejemplos de este tipo de representaciones serían yacimientos arqueológicos clasificados en función de su cronología, redes de comunicación en función de su naturaleza o lagos en función de su origen geomorfológico. Los mapas de divisiones político – administrativas también suponen un ejemplo de mapa cualitativo de polígonos.

- *Mapas cuantitativos:* son mapas que informan sobre los cambios de una variable cuantitativa, las diferencias que se establecen están determinadas por valores numéricos.

  Pueden agrupar la variable representada en intervalos o variables de resumen en función de unidades geográficas o representar la

170. Instituto Geográfico Nacional, 'Curso de Cartografía Temática. 6ª Convocatoria'.

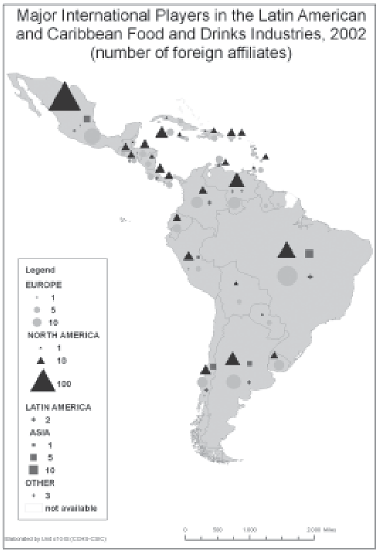


Figura 12. Representación de una variable cuantitativa en función de otra cualitativa. Fuente: Unidad SIG-CCHS (CSIC).

variación continua de una variable sobre el territorio, como sería el caso de los mapas de isolíneas. Los mapas más habituales en este apartado son los de símbolos proporcionales y los de coropletas.

Mapas de símbolos proporcionales: Son fáciles de interpretar, ya que la asociación entre cantidades y tamaños es muy intuitiva. Los símbolos utilizados suelen ser fijos (algún tipo de símbolo pictográfico o geométrico para los mapas de puntos o líneas de diferente grosor) y su tamaño variar en función de un factor de escala que debemos definir teniendo en cuenta los valores máximos y mínimos a representar o en función de la agrupación de la variable en intervalos.

En algunos casos podemos combinar la utilización de diferentes símbolos proporcionales para indicar la variación de nuestra variable cuantitativa en función de otra variable, de tipo cualitativo (Figura 12).

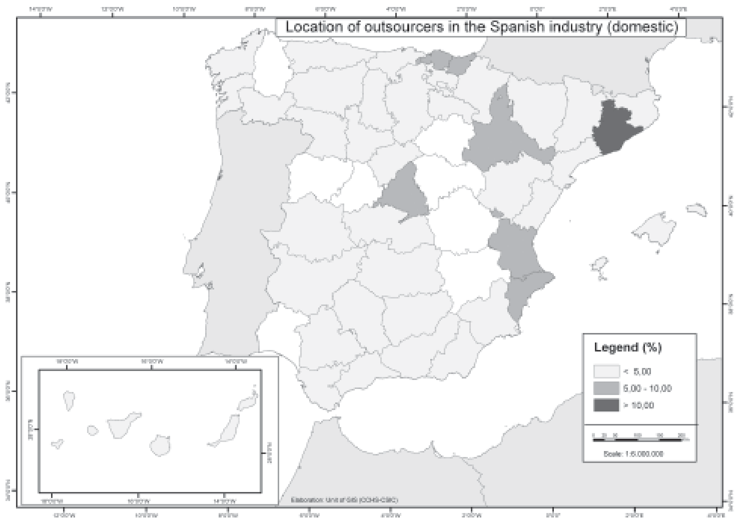


Figura 13. Industrias exportadoras agrupadas en tres intervalos.
Fuente: Unidad SIG-CCHS (CSIC)

Podemos agrupar los tipos de símbolos en lineales (barras, por ejemplo), superficiales (figuras geométricas) o volumétricos. Los lineales y superficiales son de interpretación más directa, aunque pueden ser problemáticos para representar cantidades con rangos muy amplios, mientras que los volumétricos permiten trabajar con rangos muy amplios, pero su interpretación es más problemática.

Mapas de coropletas: Cuando la información cuantitativa se asocia a áreas, representadas por polígonos, no es posible diferenciar valores a través del tamaño sin alterar la geocodificación de los polígonos. En estos casos se emplean tramas o variaciones de tono y brillo para expresar los valores numéricos. Para la elección de colores hay que tener en cuenta que generalmente asociamos los más oscuros con valores numéricos más altos.

En este tipo de representaciones cartográficas los datos se asocian a unidades territoriales, generalmente delimitaciones administrativas (ver Figura 13); cuando se distribuyen de forma continua por el territorio no son adecuados para este tipo de representaciones.

Los datos cuantitativos pueden presentarse de forma relativa cuando relacionan dos o más variables, lo que sería el caso de porcentajes o proporciones. Cuando la relación se establece entre una variable y una superficie de referencia hablamos de densidades, como la densidad de población, medida en número de habitantes por kilómetro cuadrado.

Comunicación y visualización cartográfica

La realización de mapas está al alcance de la mano de cualquier usuario de un SIG, como hemos visto, pero no debemos olvidar unas ciertas normas y recomendaciones de estilo que nos ayudarán a satisfacer las necesidades del usuario final, elaborando un mapa fácil de usar, claro, legible, atractivo y, a la vez, preciso.

Hay una serie de elementos cuya presencia no debemos descuidar. Estos son el marco de referencia con el sistema de coordenadas, indicación del Norte, la escala numérica y gráfica, el mapa de localización, así como la proyección utilizada. Los rasgos cartografiados han de ser seleccionados con cuidado en función de la escala y del objetivo del propio mapa: debemos mostrar sólo los elementos necesarios, eliminando todo lo que introduzca ruido y dificulte la lectura.

A la hora de diseñar un mapa podemos jugar con siete *variables gráficas*: tono, brillo, tamaño, forma, textura, orientación y localización. Las combinaciones de estas variables nos permiten expresar de forma gráfica el contenido temático de nuestro mapa. Así, el brillo o el tamaño son las más utilizadas para representar variaciones ordinales o numéricas, el tono (color) se suele utilizar para representar información cualitativa, la forma puede sugerir el fenómeno representado, etc. El empleo de estas variables debe ser medido para no producir un efecto de excesivo abigarramiento que dificulte la lectura y comprensión del mapa.

Nuevas herramientas para la visualización cartográfica

Aunque los mapas temáticos son posiblemente la forma más habitual de representación cartográfica en las Ciencias Humanas y So-

ciales, no podemos dejar de lado otros tipos de representaciones que están adquiriendo cada vez más importancia debido en gran medida a la disponibilidad de herramientas informáticas que permiten interactuar fácilmente con los elementos gráficos que componen la cartografía.

Mientras que los mapas temáticos proveen una tradicional representación estática, aún en gran vigencia por su concisión y elegancia —recurso habitual para la ilustración de artículos científicos y libros—, las aplicaciones informáticas ofrecen un universo en el que el usuario puede interactuar con el mapa, identificando directamente la información que le interesa. El despliegue de estos sistemas en la World Wide Web presenta características propias cuya exposición ocupa un lugar preferente en el próximo capítulo. De momento procederemos a examinar salidas alternativas a los mapas temáticos en un entorno de escritorio.

La mayor parte de los paquetes de software ofrecen herramientas que pueden mejorar ostensiblemente nuestra percepción de ciertos fenómenos geográficos:

Visualización 3D: hoy en día la posibilidad de crear visualizaciones dinámicas en 3D está abierta a un gran número de usuarios. La visualización de información geográfica supone cartografiar las dimensiones seleccionadas de un conjunto de datos, y la simplificación y abstracción de información a través de la utilización de un conjunto de variables visuales[171], como vimos anteriormente. La utilización de elementos como la perspectiva o la oclusión añaden un grado de libertad a la representación visual, reconstruyendo escenas tridimensionales sobre una proyección en dos dimensiones.

La ventaja más obvia es la representación de datos con tres dimensiones espaciales, como pueda ser el terreno, datos geológicos o atmosféricos, aunque estas representaciones no se circunscriben a datos tridimensionales; también se utiliza la tercera dimensión para represen-

171. Jo Wood et al., 'Using 3D in visualization', en *Exploring Geovisualization*, ed por Jason Dykes, Alan M. MacEachren y Menno-Jan KraakOxford: International Cartographic Association. Elsevier, 2005, pp. (295-312).

tar el valor de variables como la densidad de población, la temperatura o cualquier otra que se distribuya espacialmente.

Además de las cuestiones puramente geométricas, es necesario considerar otros aspectos a la hora de trabajar en un entorno tridimensional, como son el uso de color, sombreado, iluminación y texturas. Su utilización permite obtener entornos muy realistas superponiendo por ejemplo imágenes digitales sobre una superficie tridimensional (caso de las texturas).

En ámbitos de las Ciencias Humanas y Sociales como pueda ser la Arqueología, la modelización y visualización en 3D constituye uno de los aspectos más importantes dentro de la utilización de nuevas tecnologías. Su uso se ha extendido principalmente encaminado a la reproducción y reconstrucción de objetos arqueológicos o contextos más amplios, generalmente a través de escáner láser.

El escáner láser se ha convertido en una herramienta de trabajo de gran aplicación en la creación de representaciones exactas de objetos o contextos arqueológicos más amplios como puedan ser los hipogeos[172], permitiendo además evaluar su estado de conservación. Dentro del ámbito de la Arqueología también facilita el reconocimiento aéreo de estructuras arqueológicas[173], superando los inconvenientes de los tradicionales métodos fotográficos.

Globos virtuales: recientemente se han popularizado este tipo de aplicaciones, que permiten una visualización tridimensional de todo el planeta (algunas de las aplicaciones existentes permiten también la visualización de otros planetas).

Se trata de herramientas generalmente libres (sin coste), que permiten observar el globo terráqueo desde el espacio y acercarse hasta

172. Daniel Blersch, Marcello Balzani y Gennaro Tampone, 'The Volumni's Hypogeum in Perugia, Italy. Application of 3D survey and modelling in archaeological sites for the analysis of deviances and deformations'presentado en 2nd International Workshop, Rome: BAR International Series, 2006, mdlxviii, pp. (389-394).

173. Michael Doneus y Christian Briese, 'Full-waveform airborne laser scanning as a tool for archaeological reconnaissance'presentado en 2nd International Workshop, Rome: BAR International Series, 2006, mdlxviii, pp. (99-105).

alcanzar una resolución submétrica, soportados a través de internet y enriquecidos con todo tipo de información georreferenciada. Permiten al usuario moverse libremente en un entorno virtual tridimensional, cambiando el ángulo y la posición.

Si bien se trata de una idea desarrollada en primer lugar en el marco de enciclopedias virtuales, los primeros globos *online* fueron el NASA World Wind -2004- y el Google Earth -2005-[174]. Actualmente la oferta de globos *online* es bastante amplia, algunos de los más conocidos son, aparte de los mencionados, SkylineGlobe, CitySurf Globe, Marble (KDE) o ArcGIS Explorer. Algunos casos como World Wind o Marble son de código abierto.

Las funcionalidades básicas son las mismas, centradas generalmente en la navegación y geolocalización de diferentes tipos de información. Su principal activo es el fácil y rápido manejo, además del atractivo visual. En general no cuentan con funcionalidades analíticas al nivel de un SIG, no es ése su propósito, excepto en algún caso como SkylineGlobe o ArcGIS Explorer que, siendo un producto de la empresa ESRI, permite incorporar herramientas analíticas de su software. En palabras del experto en SIG M. Goodchild "*Just as the PC democratized computing, so systems like Google Earth will democratize GIS*"[175].

Desde nuestro punto de vista, los globos nos permiten compartir datos geográficos de un modo atractivo y sencillo. Sobre estas aplicaciones es posible mostrar nuestros datos vectoriales o georreferenciar de modo sencillo imágenes ráster. El principal inconveniente que presentaba esta práctica era el de los formatos soportados por cada navegador aunque, en un momento en el que la información geográfica tiende hacia la adopción de estándares del Open Geospatial Consortium (ver siguiente capítulo), *kml*, el formato nativo de Google Earth ha pasado a ser estándar a partir de su versión 2.2[176]. De hecho *kml* está soportado

174. 'Virtual globe - Wikipedia, the free encyclopedia', 2011 <http://en.wikipedia.org/wiki/Virtual_globe> [accedido 6 Abril 2011].

175. Declan Butler, 'Virtual globes: The web-wide world', *Nature*, 439 (2006) pp. (776-778).

176. 'KML | OGC(R)', 2011 <http://www.opengeospatial.org/standards/kml> [accedido 6 Abril 2011].

por la mayoría de globos, así como la especificación Web Map Service de OGC y los formatos más habituales de imágenes[177].

Sobre estos globos, además de poder buscar direcciones y, en algunos casos trazar rutas óptimas, podemos visualizar en 3D el terreno e imágenes de sensores remotos de diferente resolución según el nivel de zoom. Estas utilidades básicas pueden ser muy prácticas a la hora de familiarizarnos con nuestra zona de estudio. Además contamos con la posibilidad de añadir nuestras capas de información geográfica, con el aliciente, en el caso de la información ráster, de generar automáticamente un *drapping* (superposición) sobre el MDE.

El fácil y rápido acceso a la información geográfica que permiten está suponiendo un acercamiento de muchos proyectos dentro del ámbito de las Ciencias Humanas y Sociales a las tecnologías espaciales. Empieza a ser una realidad la utilización de estas herramientas en prospecciones arqueológicas[178] (la utilización de productos de teledetección en éstas ya es habitual) o incluso, el seguimiento y monitorización de probables expolios, posible gracias a la alta resolución de las imágenes servidas por Digital Globe[179], incorporadas actualmente a Google Earth.

Representación de la Dimensión Temporal

Las representaciones visuales tienen un efecto en la forma en que vemos e interactuamos con los datos[180], pueden facilitar diferentes lecturas sobre los mismos datos. La visualización geográfica no sólo debe ser conceptualizada como una forma de mostrar resultados, sino también de explorar nuestros datos en diferentes etapas de una investigación.

---

177. 'Product Comparison - World Wind Wiki', 2011 <http://worldwindcentral.com/wiki/Product_Comparison> [accedido 11 Abril 2011].

178. Anthony Beck, 'Antiquity: Project Gallery', 2011 <http://www.antiquity.ac.uk/projgall/beck308/> [accedido 20 Junio 2011].

179. Elisabeth C. Stone, 'Patterns of looting in southern Iraq', *Antiquity*, 82 (2008) pp. (125-138).

180. Monica Wachowicz, 'The role of geographic visualisation and knowledge discovery in spatio-temporal data modelling', en *Time in GIS: Issues in spatio-temporal modelling*Delft: Nederlandse Commissie voor Geodesie Netherlands Geodetic Commission, 2000, pp. (13-26).

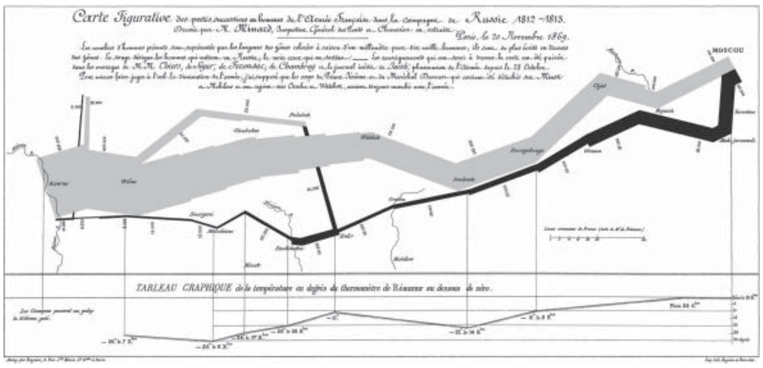


Figura 14: Campaña napoleónica en Rusia.
Fuente: Minard, 1861.

Los modelos de datos espacio-temporales suelen ser tratados de forma estática, lo que implica ciertas limitaciones, siendo la representación de esta variable bastante compleja. Representar el tiempo implica representar el cambio, en la geometría, en los atributos o en ambos[181]. Existen varias formas de abordar esta problemática:

Los *mapas estáticos sencillos*, en los que se utilizan variables gráficas y símbolos para mostrar cambio y eventos. Un claro ejemplo son los mapas históricos en los que se representa el avance de ejércitos mediante flechas y las batallas mediante espadas cruzadas. Un uso adecuado de la simbología permite representar diferentes variables y su evolución en el tiempo, como en el caso del mapa de Minard de 1861, que muestra la campaña rusa de Napoleón, con el avance y las bajas en función del grosor de la línea, en relación con la temperatura (Figura 14).

Las *series de mapas estáticos*, que representan instantes temporales. En ellas el cambio se percibe por la sucesión de mapas, que el usuario tiene que comparar visualmente. Permite la secuencia de mapas de coropletas o de cualquier otro tipo, aunque resulta problemático seguir el cambio a través de series muy largas.

---

181. Menno-Jan KRAAK, 'Visualisation of the time dimension', en *Time in GIS: Issues in spatio-temporal modelling.* Delft: Nederlandse Commissie voor Geodesie Netherlands Geodetic Commission, 2000, pp. (27-35).

Si mostramos la secuencia de mapas estáticos en una serie de pantallazos obtendremos un mapa animado en el que percibiremos el cambio como movimiento sobre el propio mapa. Las animaciones sobre las que el usuario no tiene ningún control pueden ser difíciles de entender, pero las herramientas de los SIG actuales permiten una interacción a través de la propia leyenda, en forma de barra de desplazamiento por una línea temporal, lo que favorece una correcta interpretación de los datos.

A la hora de crear mapas animados hemos de pensar no sólo en las variables gráficas (*vid. supra*), sino también en variables dinámicas como duración, orden, ritmo o frecuencia[182].Tampoco debemos olvidar la calidad de nuestros datos, que no será la misma para cada fecha, especialmente en el caso por ejemplo de SIG históricos, para lo que algunos autores proponen mostrar simultáneamente la secuencia de mapas y la información sobre la calidad.

También podemos recurrir, en el caso de cierto tipo de datos, a representaciones alternativas como pueda ser un cubo espacio-temporal en el que el eje vertical es el tiempo, para mostrar por ejemplo, rutas.

Otros tipos de representaciones

El abanico de salidas gráficas es más amplio si prescindimos de la exactitud en la geolocalización de los elementos que componen nuestro mapa.

- *Cartogramas:* son mapas en los que una variable, como pueda ser densidad de población o distancia relativa, modifica la localización de sus elementos. Los mapas de transportes urbanos como el metro son un ejemplo, ya que se busca mantener una característica principal, como es la conectividad, prescindiendo de la verdadera localización de los objetos; o mapas de coropletas en los que el tamaño de las áreas se ve modificado en función de alguna variable de interés.

182. *Ibid.*

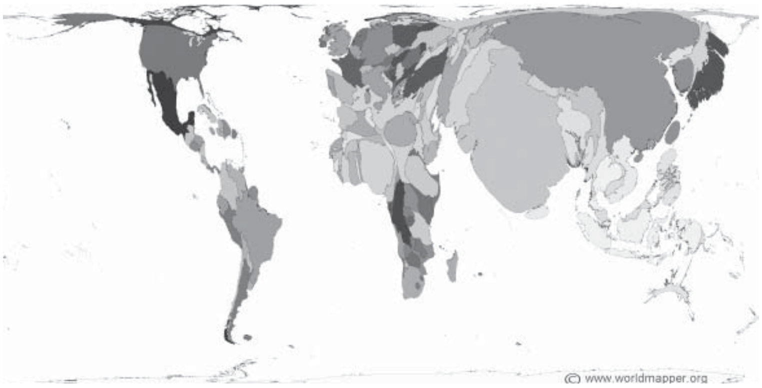


Figura 15: cartograma de población. El tamaño de los países varía en función de su población. Fuente: www.worldmapper.org. © Copyright SASI Group (University of Sheffield) and Mark Newman (University of Michigan)

Los mapas del mundo son una herramienta muy poderosa para conformar el modo en que pensamos el mundo[183], lo que hace de los cartogramas una herramienta muy eficiente para comunicar diferentes variables, especialmente usados para variables sociodemográficas (ver Figura 15).

- *Gráficos:* aún más se amplían las posibilidades de presentar información si pensamos en las diferentes salidas gráficas más allá de las estrictamente cartográficas. La primera decisión que debemos tomar cuando queremos representar los datos es si la representación va a ser sobre el espacio geográfico o el espacio definido por los propios datos[184]: podemos representar nuestros datos sobre ejes que expresen distancias o el valor de las variables sobre la superficie terrestre.

183. Anna Barford y Danny Dorling, 'Telling an Old Story with New Maps', en *Geographic Visualization. Concepts, tools and applications*, ed por Martin Dodge, Mary McDerby y Martin Turner. West Sussex: Wiley, 2008, pp. (67-107).

184. David Dibiase et al., 'Multivariate display of geographic data: applications in Earth System Science', en *Visualization in modern carthography*, ed por Alan M. MacEachren y D. R. Fraser TaylorOxford: Pergamon, 1994, ii, pp. (287-312).

Además de los clásicos gráficos bivariados o incluso tridimensionales, la combinación de los elementos espaciales y temáticos puede generar salidas gráficas útiles para el análisis paisajístico dentro de las Ciencias Humanas y Sociales:

Por ejemplo, en el estudio de uso del suelo de época romana realizado por Gugl en Panonia[185], el empleo de gráficos permite un análisis diacrónico caracterizando paisajes actuales y pretéritos a través de la orientación de sus elementos lineales. En este caso, dado que el interés del estudio se centra en la orientación, se realizaron gráficos circulares que reflejan la longitud y el acimut de estos elementos (ver Figura 16), permitiendo el análisis de diferentes hipótesis de centuriación romana.

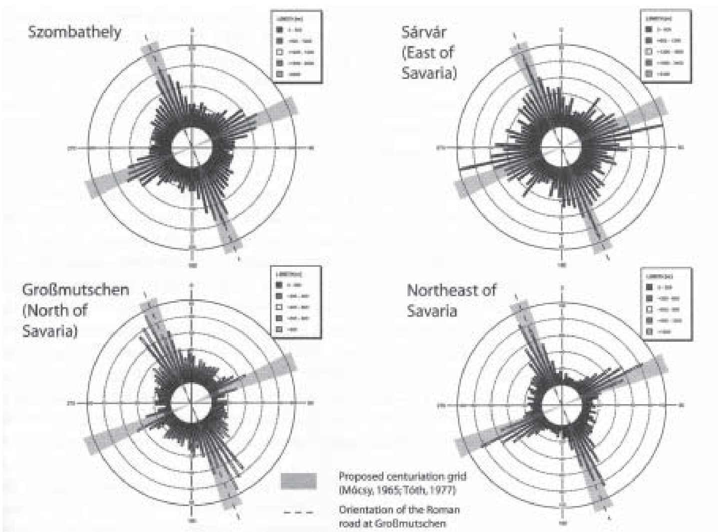


Figura 16. Gráficos circulares de orientación de elementos lineales del paisaje en la zona de Panonia[186].

185. Christian Gugl et al., 'Mapping and analysis of linear landscape features', en *Geoinformation technologies for geocultural landscapes: european perspectives.* Boca Ratón: CRC Press, 2009, pp. (275-290).

186. *Ibid.*

Capítulo VII
Difusión

El mayor peso que el pensamiento espacial está adquiriendo en las disciplinas humanas y sociales se ve reflejado en una divulgación de los resultados de proyectos de investigación, centrada en su componente espacial. En un entorno de rápida y amplia difusión a través de las tecnologías web, las herramientas espaciales van ganando terreno y popularidad, un hecho que se refleja en la forma de acceso a los datos científicos, a través de su referencia espacial.

El desarrollo tecnológico ha elevado las expectativas del público respecto a la posibilidad de obtener acceso inmediato a información espacial precisa y de confianza[187], un planteamiento que obtiene una clara respuesta en el campo de las Ciencias Naturales y poco a poco también en el de las Humanidades y las Ciencias Sociales, ya que el conocimiento generado a través de estos proyectos de investigación suele tener un claro referente geográfico. En algunos casos sus límites geográficos no tienen una delimitación clara, pero en otros casos la propia información que se maneja desde los proyectos es en sí misma información espacial, ya que describe rasgos del territorio.

La generación de información geográfica es un proceso costoso, realizado habitualmente para algún proyecto concreto. Es por ello que hoy en día se ha convertido en un tópico la necesidad de compartir y reutilizar la información geográfica, siendo la World Wide Web

187. Niels Manen et al., 'Synthesis: Geospatial Technology and the Role of Location in Science', en *Geospatial Technology and the Role of Location in Science*, Springer Netherlands. Dordrecht, 2009, pp. (303-312).

el vehículo más adecuado para poner en relación productores y consumidores de datos.

La publicación de datos a través de la red necesita cumplir con una serie de requisitos para permitir el descubrimiento, en primer lugar, de esos datos, y el acceso a los mismos a través de un servicio de descarga o consulta. Tan importantes son por lo tanto los *metadatos* como los propios *servicios web*:

Metadatos: son información sobre datos, servicios o sistemas que interactúan con las bases de datos[188]. Los metadatos permiten al usuario encontrar y evaluar la adecuación de los datos que busca para su propósito. Para cumplir con su función los metadatos deben contar al menos, con información sobre el productor de datos, la fecha de producción o modificación, la calidad y el modo de obtener los datos.

A la hora de crear, almacenar y servir metadatos es fundamental seguir las directrices marcadas por los estándares más comunes, CSDGM (FGDC), Z39.50 GEO Profile o ISO 19115, para permitir la interoperabilidad entre catálogos de metadatos, de forma que un usuario pueda buscar simultáneamente los datos que le interesan en varios servidores.

Servicios web: existen dos formas de acceder a datos de nuestro interés, la descarga para una posterior elaboración en local, o el acceso a servicios externos, servicios que se pueden consumir desde un cliente pesado (software de SIG instalado en local) o ligero (aplicaciones desplegadas en navegadores web).

Las administraciones nacionales disponen de centros de descarga de datos geográficos, como en el caso español del Centro Nacional de Información Geográfica (CNIG)[189], que ofrece toda la información geográfica digital generada por el Instituto Geográfico Nacional (IGN) a partir de una orden ministerial de 2008 (FOM/956/2008).

188. Albert K. W. Yeung y G. Brent Hall, *op. cit.,* p. (155).
189. http://centrodedescargas.cnig.es/CentroDescargas/index.jsp.

Sin embargo, hoy en día se tiende más a ofrecer servicios que datos. Los servicios se basan en la arquitectura cliente–servidor, constituida por clientes que solicitan servicios y servidores que responden a sus peticiones. Desde la parte servidor se ofrecen funcionalidades que van desde la visualización más sencilla de un mapa hasta diferentes tipos de geoprocesamiento o análisis espacial.

Compartir la Información Geográfica

El acceso a información espacial presenta problemas de disponibilidad, calidad, organización, accesibilidad y puesta en común, abordados en el seno de la Unión Europea desde la directiva INSPIRE (2007/2/CE), que propugna el establecimiento de una infraestructura de información espacial comunitaria. El marco de trabajo es el establecimiento de servicios de red, que deben garantizar la interoperabilidad y hacer posible *localizar, transformar, visualizar y descargar datos espaciales*.

Dentro de este marco es fundamental el establecimiento de Infraestructuras de Datos Espaciales (IDE), como vehículo para la difusión de datos y conocimiento. Podemos definir una IDE como un sistema informático dedicado a gestionar información geográfica, disponible en Internet, que cumple una serie de condiciones de interoperabilidad (normas, especificaciones, protocolos, interfaces…) que permiten que un usuario, utilizando un simple navegador, pueda utilizarlos y combinarlos según sus necesidades[190].

El desarrollo de las IDE requiere la estandarización de datos y de servicios. Respecto a los datos, actualmente se están desarrollando especificaciones para todos los tipos de datos incluidos en la directiva, teniendo siempre como base la adaptación a la serie ISO19100, el conjunto de normas internacionales referentes a la información geográfica. Las especificaciones de servicios se definen desde el Open Geospatial Consortium (OGC), siendo el Web Map Service (WMS)

190. 'Portal IDEE', 2011 <http://www.idee.es/show.do?to=pideep_que_es_IDEE.ES> [accedido 22 Agosto 2011].

la especificación básica más utilizada, consistente en servir mapas a demanda en formato de imagen, con la posibilidad de identificar elementos. Aunque éste es el servicio básico más extendido, existen muchas otras especificaciones que facilitan la realización de una gran cantidad de operaciones: Web Feature Service (WFS), para el acceso a capas vectoriales de información, Web Coverage Service (WCS), para datos ráster, Catalog Service Web (CSW) para la consulta de catálogos de metadatos, etc.

La directiva INSPIRE está encaminada a la gestión e implementación de políticas sobre el territorio comunitario con la protección como primer objetivo claro. Aunque la preocupación principal de la directiva es el medio ambiente, dentro de los lugares protegidos está incluido el Patrimonio Cultural, lo cual ha despertado el interés de investigadores de este ámbito.

Por otro lado, conviene resaltar el espíritu de la declaración Open Access de Berlín[191] firmada por 302 organizaciones científicas, entre ellas el CSIC: *por primera vez Internet ofrece la posibilidad de constituir una representación global e interactiva del conocimiento humano [...] las contribuciones de libre acceso incluyen resultados originales de investigación, datos primarios y metadatos...*

Dentro de este contexto muchos proyectos desarrollados por esta institución buscan en el ámbito de los visores geográficos una forma sencilla e intuitiva de conectar a los usuarios con los resultados de la investigación científica. Como ya vimos en la introducción, la acogida desde el CCHS ha sido bastante significativa con la aparición de varias IDE y un creciente interés de todos los proyectos con una clara componente espacial por contribuir a la difusión del conocimiento generado a través de IDE.

Desde el ámbito de la Arqueología se ha visto desde muy temprano[192] la necesidad de trabajar en línea con la estandarización y

191. 'OA MPG»Berlin Declaration', 2011 <http://oa.mpg.de/berlin-prozess/berliner-erklarung/> [accedido 10 Agosto 2011].

192. Daniela Ballari, Miguel Ángel Manso Callejo y Miguel Ángel Bernabé Poveda, 'Arqueología y Servidores de Mapas en Red. Proyecto LIFE «Valle De Tiermes – Ca-

la creación de metadatos, necesarios para asegurar la interoperabilidad demandada desde INSPIRE. En el caso del CCHS, se está abordando la difusión de los datos de proyectos arqueológicos a través de una IDE que cumpla con las especificaciones OGC y de INSPIRE, lo que implica un gran esfuerzo en cuanto a estandarización de datos y servicios. En esta misma línea, del mismo centro son otras iniciativas como la creación de una IDE para gestionar los recursos sociales para personas mayores en España, la IDE de cartografía y demografía histórica de la ciudad de Madrid o la futura IDE lingüística en base al proyecto ALPI, entre otros.

Interacción en la Red

Este desarrollo está alimentado también por la rápida adopción de prácticas de la Web 2.0 como el geoetiquetado de información o de fotografías sobre diferentes servicios, la accesibilidad de sistemas GPS o los programas y herramientas de "mapeo", que contribuyen a alimentar el interés por el pensamiento espacial dentro de las propias Ciencias Humanas y Sociales[193].

Recientes eventos de actualidad han contribuido a mostrar esta popularización de las herramientas espaciales y los beneficios que de ella se pueden esperar. Así, pudimos observar algunos aspectos de las revueltas populares de Egipto a principios de 2011 mediante el contenido volcado por miembros de las redes sociales como fotos, vídeos o mensajes, recogido *in situ* y geolocalizado en una aplicación disponible en Internet[194].

racena»' (presentado en III Jornadas Técnicas de la IDE de España, Madrid, 2005); María Aranzazu Respaldiza, 'La capa patrimonial de la IDEE. Importancia del Patrimonio Arqueológico como capa interoperable' (presentado en las I Jornadas de jóvenes en investigación arqueológica: dialogando con la Cultura Material, Madrid, 2008), pp. (123-128); María. Aranzazu Respaldiza, Antonio Vázquez Hoenhe y A. Wachowicz, 'Propuesta de un Núcleo Estándar de Metadatos para los recursos del Patrimonio Histórico Español', 2009; Anthony J. Corns y Robert Shaw, 'Cultural Heritage Spatial Data Infrastructures (SDI) - Unlocking the Potential of Our Cultural Landscape Data', en *30th EARSeL Symposium Remote Sensing for Science, Education, and Natural and Cultural Heritage*, ed Rainer Reuter, 2010, pp. (1-8).

193. Donald G. Janelle y M. F Goodchild, 'Location across Disciplines: Reflections on the CSISS Experience', en *Geospatial Technology and the Role of Location in Science*, Springer Netherlands. Dordrecht, 2009, pp. (15-30).

194. http://tmapps.esri.com/egypt_unrest/index.html.

Anteriores eventos como el huracán Katrina (2005) o el terremoto de Pakistán (2005) mostraron cómo las nuevas herramientas cartográficas, que ofrecen imágenes de alta resolución espacial al alcance de un amplio público, pueden crear un nuevo "cuerpo" de voluntarios, relevante para la organización de operaciones de salvamento[195]. El caso más evidente ha sido el reciente terremoto de Haití, en el que, ante la destrucción de la única agencia cartográfica haitiana los propios voluntarios de OpenStreetMap y Google MapMaker realizaron un levantamiento cartográfico de las principales ciudades, necesario para la coordinación de las labores de ayuda[196].

En todo caso, la generalización de la denominada Web 2.0 ha cambiado el patrón de interacción en la red: los usuarios añaden valor, siendo también productores de datos, y la cooperación se impone al control. Estos conceptos ya se están integrando en las IDE[197], e incluso en la generación de cartografía[198], y pueden enseñar valiosas lecciones al desarrollo de proyectos en el ámbito de las Ciencias Sociales y las Humanidades[199], como la importancia de la referencia espacial a modo de contexto natural para la integración de información interdisciplinar, y como nexo de teoría y datos.

En un entorno en el que la calidad de la información es crucial, puede resultar problemático confiar en la denominada "Información Geográfica Voluntaria", proporcionada por usuarios no expertos en la materia, pero el gran potencial de la Red puede ser positivo estableciendo los cimientos para una generación de "Información Geográfica Voluntaria Acreditada"[200].

---

195. Iván Sánchez Ortega, 'El papel de las IDEs en el año que cambió la ayuda humanitaria' presentado en las I Jornadas ibéricas de Infra-estructuras de Dados Espaciais, Lisboa, 2010.

196. Silvia Laiglesia et al., 'Hacia una IDEE 2.0: integrando a los usuarios y sus contenidos' presentado en las VI Jornadas Técnicas de la IDE de España JIDEE2009, Murcia, 2009.

197. Illah Nourbakhsh et al., 'Mapping disaster zones', *Nature*, 439 (2006) pp. (787-788).

198. 'OpenStreetMap', 2011 <http://www.openstreetmap.org/> [accedido 23 Agosto 2011].

199. Donald G. Janelle y M. F Goodchild, *op. cit.,* pp. (15-30).

200. Ernest Ruiz i Almar, 'Consideraciones acerca de la explosión geográfica: geografía colaborativa e información geográfica voluntaria acreditada.', *Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 10 (2010) pp. (280-298).

Desde centros pioneros en la aplicación de los SIG en Ciencias Sociales como el CSISS (Center for Spatially Integrated Social Sciences) se señala la oportunidad de las circunstancias actuales para incorporar al mundo de las Ciencias Sociales esta corriente[201].

201. Donald G. Janelle y M. F Goodchild, *op. cit.,* p. 23.

Conclusiones

Es un hecho innegable el mayor peso que el pensamiento espacial está adquiriendo en las investigaciones en el campo de las Ciencias Humanas y Sociales: Cada vez son más los proyectos que incorporan análisis espaciales y por ende algún tipo de Tecnología de Información Geográfica en el marco de la investigación, un hecho posiblemente relacionado con su potencial como elemento integrador de diferentes disciplinas sobre un espacio común a todas ellas: el territorio.

A pesar de esta mayor influencia del pensamiento geográfico en las Ciencias Humanas y Sociales, la difusión varía notablemente de unas disciplinas a otras. La Arqueología y la Historia siempre han necesitado desplegar sus argumentos sobre un contexto geográfico, pero desde los años 60', parte de la Arqueología se ha orientado fuertemente hacia problemas de localización, dándose el paso hacia la explotación de las capacidades analíticas ofrecidas por los SIG. En el campo de la Historia las posibilidades que ofrecen los SIG son similares al caso de la Arqueología, pero su aprovechamiento mucho menor, más centrado en cuestiones de representación de la información que en su análisis.

En disciplinas en las que ya existía una tradición de estudios territoriales, los SIG favorecen un desarrollo de análisis cuantitativos, como en el caso de la Demografía; en algunos campos de investigación la adopción de este tipo de tecnologías está ligada al compromiso de investigadores que pueden convertirse en un referente para una implementación más generalizada y en muchos casos es la, cada vez mayor, disponibilidad de amplios conjuntos de datos georreferenciados el acicate para la adopción de estas tecnologías.

Las Tecnologías de Información Geográfica por su parte han protagonizado un avance espectacular en los últimos años, tanto en funcionalidades como en disponibilidad y accesibilidad para nuevos usuarios que hasta hace muy poco no tenían acceso a ellas. Esto puede suponer su definitiva integración en el ámbito de las Ciencias que nos ocupan en el presente libro, en una época en la que los participantes están habituados a utilizar las funcionalidades desplegadas por los SIG en un entorno cotidiano: búsqueda de direcciones, observación remota de la superficie terrestre, sistemas de navegación, etc.

La generalización del pensamiento geográfico viene acompañada de un acceso directo a gran cantidad de información geográfica a través de la red, y a las posibilidades de análisis y representación que los SIG de escritorio nos ofrecen. Los organismos encargados de la información geográfica en cada país, así como en la Unión Europea, ofrecen de forma gratuita una gran cantidad de datos; productos disponibles, también de forma gratuita, nos han ayudado a familiarizarnos con imágenes propias del campo de la Teledetección; en el campo del análisis espacial y especialmente en el de modelos ráster, se están aplicando metodologías cada vez más avanzadas que incluyen por ejemplo novedosas técnicas para el estudio del modelado de patrones espaciales, como son el análisis a través de la geometría fractal, el modelado con estadística espacial y los autómatas celulares; las posibilidades de representación y difusión de nuestros datos se han elevado de forma exponencial gracias a las nuevas técnicas de geovisualización.

Estos elementos hacen que los SIG se vayan abriendo un espacio propio como proveedores de datos, como herramientas para el análisis, para la representación de datos o para la difusión de resultados; pero su auténtica potencialidad para las Ciencias Humanas y Sociales puede estar integrada en la propia definición de la investigación. Son los casos en los que los investigadores buscan traducir cuestiones propias de su ámbito al lenguaje propio de los modelos geográficos; en esos casos el análisis espacial aporta un marco de trabajo que facilita la generación de conocimiento científico a través de la contrastación de hipótesis, entendiendo el conocimiento científico como aquél cuyas conclusiones pueden ser discutidas independientemente de las condiciones subjetivas en que se ha generado.

Además de los avances que los SIG pueden deparar mediante su aplicación en parcelas concretas del saber, la disponibilidad de datos y servicios en la red, claramente vinculada a la progresiva implantación de INSPIRE, hace de las Infraestructuras de Datos Espaciales un objetivo cada vez más identificable, y deseable, en la concepción de los proyectos de Humanidades y Ciencias Sociales. La aplicación de los mismos criterios de estandarización e interoperabilidad a datos generados por proyectos de nuestro ámbito puede abrir el camino a un futuro de integración de líneas de investigación más allá de la interdisciplinariedad, con planteamientos y objetivos comunes.

Los estándares marcados por las organizaciones de normalización de información geográfica, como el Comité Técnico 211 (ISO/TC 211), responsable del desarrollo de la serie ISO 19100, el Open Geospatial Consortium (OGC), responsable de estándares como el Web Map Service (WMS) o incluso el World Wide Web Consortium (W3C), jalonan el camino hacia la integración de datos provenientes de diferentes disciplinas. Los estándares propios que puedan existir dentro de cada campo de aplicación no habrán de ser un obstáculo, al afectar más a la parte temática que a la geográfica, y su convivencia con los estándares de Información Geográfica sólo puede ser beneficiosa. En todo caso, en disciplinas en las que no es habitual la estandarización de datos, esta corriente integradora puede suponer un interesante punto de partida.

El propósito de esta obra es adentrarse en la teoría y los fundamentos de los Sistemas de Información Geográfica (SIG) desde la perspectiva de su implementación en las Ciencias Humanas y Sociales.

En este libro se presentan los fundamentos básicos para el diseño y el modelado de datos espaciales y la generación de bases de datos georreferenciadas, se abordan las funciones elementales de consulta y edición, análisis espacial y representación de la información mediante técnicas cartográficas, para terminar con su difusión y publicación.

Esperamos proporcionar unas pautas que permitan la comprensión de lo que es un SIG, cómo funciona y cómo su aplicabilidad técnica puede incidir favorablemente en la actividad investigadora en estas disciplinas.

CSIC
Consejo Superior de Investigaciones Científicas
CONFEDERACIÓN ESPAÑOLA DE
CENTROS DE ESTUDIOS LOCALES (CECEL)